

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:42:39 ; Search time 14 Seconds

(without alignments)
2044.190 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690

Sequence: 1 MRAAPLLIARAASLSIGFLF.....EPVOGAPPPAAHHHHH 690

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	292	42.3	1255	1	ERR2_HUMAN
2	127	18.4	144	1	CSF2_HUMAN
3	67	9.7	1254	1	ERR2_MESAU
4	52	7.5	1257	1	ERR2_RAT
5	35	5.1	386	1	PPAP_HUMAN
6	22	3.2	144	1	CSF2_SHEEP
7	20	2.9	245	1	ERR2_MOUSE
8	13	1.9	140	1	CSF2_CAVPO
9	11	1.6	144	1	CSF2_CANFA
10	11	1.6	144	1	CSF2_CEREL
11	11	1.6	144	1	CSF2_PIG
12	10	1.4	143	1	CSF2_BOVIN
13	9	1.3	530	1	ZIC2_MOUSE
14	9	1.3	532	1	ZIC2_HUMAN
15	9	1.3	703	1	EGFR_CHICK
16	9	1.3	1210	1	EGFR_HUMAN
17	9	1.3	1210	1	EGFR_MOUSE
18	8	1.2	127	1	CSF2_RAT
19	8	1.2	265	1	UCR1_SOLIU
20	8	1.2	325	1	MODD_MYCHO
21	8	1.2	325	1	MODD_MYCTU
22	8	1.2	355	1	MONA_ARTNI
23	8	1.2	428	1	FXB2_MOUSE
24	8	1.2	463	1	VY30_MYCTU
25	8	1.2	547	1	CITR_KLEBN
26	8	1.2	715	1	CLPB_MYCPN
27	8	1.2	825	1	ICP0_HSV2H
28	8	1.2	1848	1	CCPA_DROME
29	7	1.0	62	1	SECE_SULSO
30	7	1.0	68	1	RI29_ARCFU
31	7	1.0	104	1	YIF4_ARCTU
32	7	1.0	111	1	MADE_HUMAN
33	7	1.0	125	1	CALR_CHICK

ALIGNMENTS

RESULT 1	ID	ERR2_HUMAN	STANDARD:	PRT: 1255 AA.
AC	P04626:			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).			
DE	surface receptor HER2 (MLN 19).			
GN	ERRB2 OR HER2 OR NGL OR NEU.			
OC	Human sapliens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86118653; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,			
RA	Salto T., Toyoshima K.;			
RT	"Similarity of protein encoded by the human c-erbB-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86070181; PubMed=2999974;			
RA	Consensus L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,			
RA	Francine U., Levinson A., Ullrich A.;			
RT	Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
RN	[3]			
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=2995967;			
RA	Semba K., Kamata N., Toyoshima K., Yamamoto T.;			
RA	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a			
RT	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
RN	[4]			
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,			
CC	ALTHOUGHOUT NEUREGULINS DO NOT INTERACT WITH IT ALONE. GTF30 IS A			
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-			
CC	ALPHA AND AMPHIREGULIN.			
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein			
CC	tyrosine phosphate.			
CC	- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS			
CC	(POTENTIAL).			
CC	- SUBCELLULAR LOCATION: Type I membrane protein.			

34	7	1.0	136	1	CAL1_ONCKE	P01263 oncorhynchu
35	7	1.0	138	1	CAL_CHICK	P07666 gallus galli
36	7	1.0	141	1	CSF2_MOUSE	P01587 mus musculus
37	7	1.0	144	1	CSF2_FELCA	O62757 felis silve
38	7	1.0	182	1	NCC2_RHME	P56896 rhizobium m
39	7	1.0	184	1	RUS_THEMA	P78517 thermotoga
40	7	1.0	186	1	RK12_NICSY	P36688 nicotiana s
41	7	1.0	186	1	RK12_TOBAC	I24929 nicotiana t
42	7	1.0	192	1	UREE_ALCEU	C70338 alcaligenes
43	7	1.0	196	1	GIDB_AQUAE	O67522 aquifex aeo
44	7	1.0	204	1	GIDB_COXBU	P94614 coxiella bu
45	7	1.0	204	1	HAN1_SHEEP	O28555 ovis aries

FT	DISULFID	563	576	BY SIMILARITY.
FT	DISULFID	567	584	BY SIMILARITY.
FT	DISULFID	587	596	BY SIMILARITY.
FT	DISULFID	600	623	BY SIMILARITY.
FT	DISULFID	626	634	BY SIMILARITY.
FT	DISULFID	630	642	BY SIMILARITY.
FT	MOD_RES	1139	1139	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	MOD_RES	1248	1248	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	CARBOHYD	68	68	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	124	124	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	187	187	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	259	259	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	530	530	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	571	571	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	629	629	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	654	654	I -> V.
FT	VARIANT	655	655	/FtId=VAR_004077.
FT	VARIANT	655	655	I -> V.
FT	CONFLICT	1170	1170	/FtId=VAR_004078.
FT	SEQUENCE	1255	137909	P -> A (IN REF. 2).
SO	SEQUENCE	1255	137909	MM; 39E9DEDA04DCE962 CRC64;
Query Match				
Best Local Similarity 100.0%; Pred. No. 78-274; Length 1255;				
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	38	GAATGCTGTGDMKRLRLPASPETHLMDLRHLGYCGQVQVGNLELTYLPTNLSFLDIO	97	
Db	19	GAATGCTGTGDMKRLRLPASPETHLMDLRHLGYCGQVQVGNLELTYLPTNLSFLDIO	78	
QY	98	EVQGVGLIHHNOVQVPLQRLRIYRGQLFEDNALAVLDGDLNNTPTVTGASPGGLR	157	
Db	79	EVQGVGLIHHNOVQVPLQRLRIYRGQLFEDNALAVLDGDLNNTPTVTGASPGGLR	138	
QY	158	ELQRLSLTEILKGVGLIQRNPQLCYODILKMDIFHKNNQALILIDTNSRACHPCSPM	217	
Db	139	ELQRLSLTEILKGVGLIQRNPQLCYODILKMDIFHKNNQALILIDTNSRACHPCSPM	198	
QY	218	CKSGRCWGESSEDCQSILRTVYAGGACARCKRPLTDCDCEGCAAGCTGPKHSIDLCLHF	277	
Db	199	CKSGRCWGESSEDCQSILRTVYAGGACARCKRPLTDCDCEGCAAGCTGPKHSIDLCLHF	258	
QY	278	NHSGICELHCAALVTYNTDTPESMPNBEGRTPFASCVTACPRYKLTDSGS	329	
Db	259	NHSGICELHCAALVTYNTDTPESMPNBEGRTPFASCVTACPRYKLTDSGS	310	
RESULT 2				
CSF2_HUMAN				
ID	CSF2_HUMAN	STANDARD;	PRT;	144 AA.
AC	P04141;			
DT	01-NOV-1986 (Rel. 03, Created)			
DT	01-NOV-1986 (Rel. 03, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)			
DE	(Colony-stimulating factor) (CSF) (Sargramostin) (Molgramostin).			
GN	CSF2 OR GMCSF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85242684; PubMed=3925454;			
RA	Lee F., Yokota T., Otsuka T., Gemmell L., Larson N., Luh J.,			
RA	Arai K.-I., Rennick D.;			
RT	"Isolation of cDNA for a human granulocyte-macrophage			
RT	colony-stimulating factor by functional expression in mammalian			
RT	cells."			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:4360-4364(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86205844; PubMed=3486413;			

RA Kaushansky K., O'Hara P.J., Berkner K., Segal G.M., Hagen F.S.,
 RA Adamson J.W.;
 RT "Genomic cloning, characterization, and multilineage growth-promoting
 RT activity of human granulocyte-macrophage colony-stimulating factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3101-3105(1986).
 RN [13]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85298329; PubMed=3898082;
 RA Cantrell M.A., Anderson D., Cerretti D.P., Price V., McKereghan K.,
 RA Tushinski R.J., Mochizuki D.Y., Larsen A., Grabstein K., Cosman D.;
 RT "Cloning, sequence, and expression of a human granulocyte/macrophage
 RT colony-stimulating factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6250-6254(1985).
 RN [14]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85218749; PubMed=3923623;
 RA Wong G.G., Wittek J.S., Temple P.A., Wilkens K.M., Leary A.C.,
 RA Luxenberg D.P., Jones S.S., Brown E.L., Kay R.M., Orr E.C.,
 RA Shoemaker C., Golde D.W., Kaufman R.J., Hewick R.M., Wang E.A.,
 RA Clark S.C.;
 RT "Human GM-CSF: molecular cloning of the complementary DNA and
 RT purification of the natural and recombinant proteins.";
 RL Science 228:810-815(1985).
 RN [15]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86030234; PubMed=3876930;
 RA Miyake S., Otsuka T., Yokota T., Lee F., Araki K.-I.;
 RT "Structure of the chromosomal gene for granulocyte-macrophage colony
 RT stimulating factor: comparison of the mouse and human genes.";
 RL EMBO J. 4:2561-2568(1985).
 RN [16]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92144609; PubMed=1737041;
 RA Kaushansky K., Lopez J.A., Brown C.B.;
 RT "Role of carbohydrate modification in the production and secretion of
 RT human granulocyte macrophage colony-stimulating factor in genetically
 RT engineered and normal mesenchymal cells.";
 RL Biochemistry 31:1881-1886(1992).
 RN [19]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=92108420; PubMed=1837174;
 RA Diederichs K., Boone T., Karpus P.A.;
 RT "Novel fold and putative receptor binding site of
 RT granulocyte-macrophage colony-stimulating factor.";
 RL Science 254:1779-1782(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=92235844; PubMed=1569568;
 RA Walter M.R., Cook W.J., Ealick S.E., Nagabhushan T.L., Trotta P.P.,
 RA Bugg C.E.;
 RT "Three-dimensional structure of recombinant human granulocyte-
 RT macrophage colony-stimulating factor.";
 RL J. Mol. Biol. 224:1075-1085(1992).
 CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
 CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
 CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES.
 CC -1- SUBUNIT: MONOMER.
 CC -1- POLYMORPHISM: VARIANT ILE-117 MAY BE A RISK FACTOR FOR APOPTIC
 CC ASPHMA.
 CC -1- PHARMACEUTICAL: Available under the names Leukine (Immunex) and
 CC Leucomax (Novartis). Used in myeloid reconstitution following bone
 CC marrow transplant, bone marrow transplant engraftment failure or

CC delay, mobilization and following transplantation of autologous
 CC peripheral blood progenitor cells, and following induction
 CC chemotherapy in older adults with acute myelogenous leukemia.
 CC -1- DATABASE: NAME=Leukine; NOTE=Clinical information on Leukine;
 CC WWW="http://www.immunex.com/patient/pa02el.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M13207; AA98768.1; -
 DR EMBL: M11734; AA52122.1; -
 DR EMBL: M11220; AA52578.1; -
 DR EMBL: X03021; AA26822.1; -
 DR EMBL: M10663; AA52121.1; -
 DR EMBL: AC004511; AAC08707.1; -
 DR EMBL: AF373868; AA51563.1; -
 DR PIR: A01853; FOHGM.
 DR PIR: G24636; C24636.
 DR PIR: A25169; A25169.
 DR PDB: 1CSG; 31-JAN-94.
 DR PDB: 2GMF; 08-NOV-96.
 DR Genew; HGNC:2434; CSF2.
 DR MIM: 138960; -
 DR InterPro; IPR000773; GM-CSF.
 DR Pfam; PF01109; GM-CSF; 1.
 DR PRINTS; PR00693; GMCSFACTOR.
 DR ProDom; PD007349; GM-CSF; 1.
 DR SMART; SM00040; CSF2; 1.
 DR PROSITE; PS00702; GM-CSF; 1.
 KW Cytokine; Growth factor; Glycoprotein; Signal; 3D-structure;
 KW Polymorphism; Pharmaceutical.
 FT SIGNAL 1 17
 FT CHAIN 18 144
 FT
 FT DISULFID 71 113
 FT DISULFID 105 138
 FT CARBOHYD 22 22
 FT CARBOHYD 24 24
 FT CARBOHYD 26 26
 FT CARBOHYD 27 27
 FT CARBOHYD 44 44
 FT CARBOHYD 54 54
 FT VARIANT 115 115
 FT
 FT VARIANT 117 117
 FT
 FT TURN 25 27
 FT HELIX 30 44
 FT TURN 45 45
 FT TURN 50 54
 FT STRAND 56 60
 FT HELIX 72 81
 FT TURN 82 82
 FT HELIX 85 103
 FT TURN 104 104
 FT STRAND 115 119
 FT HELIX 120 131
 FT TURN 132 133
 FT TURN 133 133
 SO SEQUENCE 144 AA; 16225 MW; 75D1E50506BCA7A8 CRC64;
 QY Query Match 18.4%; Score 127; DB 1; Length 144;
 DB Best Local Similarity 100.0%; Pred. No. 4,2e-115;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 549 AARSPSPSTQWEHVNATQEARLLNLSRDPAENKNEVEYISEMFDLOEPTCLOTRLE 608
 DB 18 AARSPSPSTQWEHVNATQEARLLNLSRDPAENKNEVEYISEMFDLOEPTCLOTRLE 77

```

Oy 609 LYKGLRGLSTFKLKGPLTMASHYKQHCPPNPETSCATQITFESEKNTKDFLIVIPD 668
Db 78 LYKGLRGLSTFKLKGPLTMASHYKQHCPPNPETSCATQITFESEKNTKDFLIVIPD 137
Oy 669 CMEPVQ 675
Db 138 CMEPVQ 144

RESULT 3
ERB2_MESAU STANDARD; PRT; 1254 AA.
ID ERB2_MESAU STANDARD; PRT; 1254 AA.
AC 060533;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
GN ERB2 OR NEU.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Nerve;
RX MEDLINE=94193007; Pubmed=7908275;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
RA Yamazaki Y., Ishikawa T.;
RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";
RL Gene 140:251-255(1994).
CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D16295; BAA03801.1; -
DR HSSP; P11362; IFGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_Pkinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam; PF00069; pkinase.1.
DR Pfam; PF00757; Furin-like.1.
DR Pfam; PF01030; Recept_L_domain.2.
DR Pfam; PF02757; YLP.2.
DR ProDom; PD000001; Euk_Pkinase.1.
DR SMART; SM00261; FYKc.1.
DR SMART; SM00219; TYKc.3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR.1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM.1.
DR Transmembrane; Glycoprotein; MultiGene family; Receptor; Signal;
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW Proto-oncogene; Disease mutation.

```

```

FT SIGNAL 1 21
FT CHAIN 22 1254
FT DOMAIN 22 652
FT TRANSMEM 653 675
FT DOMAIN 676 1254
FT DOMAIN 158 368
FT DOMAIN 472 644
FT DOMAIN 720 987
FT NP_BIND 726 734
FT BINDING 753 753
FT ACT_SITE 845 845
FT DISULFID 195 204
FT DISULFID 199 212
FT DISULFID 236 244
FT DISULFID 240 252
FT DISULFID 255 264
FT DISULFID 268 295
FT DISULFID 299 311
FT DISULFID 315 331
FT DISULFID 334 338
FT DISULFID 511 520
FT DISULFID 515 528
FT DISULFID 531 540
FT DISULFID 544 560
FT DISULFID 563 576
FT DISULFID 567 584
FT DISULFID 587 596
FT DISULFID 600 623
FT DISULFID 626 634
FT DISULFID 630 642
FT MOD_RES 1139 1139
FT MOD_RES 1247 1247
FT CARBOHYD 68 68
FT CARBOHYD 125 125
FT CARBOHYD 187 187
FT CARBOHYD 259 259
FT CARBOHYD 530 530
FT CARBOHYD 571 571
FT CARBOHYD 629 629
FT VARIANT 658 658
FT VARIANT 659 659
SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F1B81 CRC64;

Query Match 9.78; Score 67; DB 1; Length 1254;
Best Local Similarity 100.0%; Pred. No. 2,8e-56;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 250 LPTCCHEQCAAGCTGPKHSDCLACLPNHSIGICELHCPALVTYNTDFESMPNDEGRYT 309
Db 231 LPTCCHEQCAAGCTGPKHSDCLACLPNHSIGICELHCPALVTYNTDFESMPNDEGRYT 290
Oy 310 FGASCVT 316
Db 291 FGASCVT 297

RESULT 4
ERB2_RAT STANDARD; PRT; 1257 AA.
ID ERB2_RAT STANDARD; PRT; 1257 AA.
AC P06494;
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Epidermal growth factor
DE receptor-related protein).
GN ERBB2 OR NEU.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

CC TISSUE-Neuroblastoma;
 CC MEDLINE-86118662; PubMed-3945311;
 RA Bargmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related
 RT protein.";
 RT Nature 319:226-230(1986).
 RL (2)
 RN SEQUENCE OF 852-905 FROM N.A.
 RP TISSUE-Sciatic nerve;
 RC MEDLINE-91222560; PubMed-2025425;
 RX Lal C., Lemke G.;
 RA "An extended family of protein-tyrosine kinase genes differentially
 RT expressed in the vertebrate nervous system.";
 RT Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668;
 RA MEDLINE-92155181; PubMed-1346763;
 RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-
 RT oncogenic and oncogenic forms of the neu protein.";
 RL EMO J. 11:43-48(1992).
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX.
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND AMPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER EBBB RECEPTORS.
 CC THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY:
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X03362; CAA27059.1; ALT_INIT.
 DR PIR: A24562; TVRTNU.
 DR HSSP: P11362; 1FGC.
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L.domain; 2.
 DR Pfam: PF02737; YLP; 2.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR SMART: SMO0261; Fy; 3.
 DR SMART: SMO0219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE EBBB-2.
 FT DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 655 677 POTENTIAL.
 FT DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 159 369 CYS-RICH.
 FT DOMAIN 473 646 CYS-RICH.
 FT DOMAIN 722 989 PROTEIN KINASE.
 FT NP_BIND 728 736 ATP (BY SIMILARITY).

FT BINDING 755 755 ATP (BY SIMILARITY).
 FT ACN SITE 847 847 BY SIMILARITY.
 FT DISULFD 196 205 BY SIMILARITY.
 FT DISULFD 200 213 BY SIMILARITY.
 FT DISULFD 221 228 BY SIMILARITY.
 FT DISULFD 225 236 BY SIMILARITY.
 FT DISULFD 237 245 BY SIMILARITY.
 FT DISULFD 241 253 BY SIMILARITY.
 FT DISULFD 256 265 BY SIMILARITY.
 FT DISULFD 269 296 BY SIMILARITY.
 FT DISULFD 300 312 BY SIMILARITY.
 FT DISULFD 316 332 BY SIMILARITY.
 FT DISULFD 335 339 BY SIMILARITY.
 FT DISULFD 513 522 BY SIMILARITY.
 FT DISULFD 517 530 BY SIMILARITY.
 FT DISULFD 533 542 BY SIMILARITY.
 FT DISULFD 546 562 BY SIMILARITY.
 FT DISULFD 565 578 BY SIMILARITY.
 FT DISULFD 569 586 BY SIMILARITY.
 FT DISULFD 589 598 BY SIMILARITY.
 FT DISULFD 602 625 BY SIMILARITY.
 FT DISULFD 628 636 BY SIMILARITY.
 FT DISULFD 632 644 BY SIMILARITY.
 FT MOD_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;
 Query Match 7.54; Score 52; DB 1; Length 1257;
 Best Local Similarity 100.00; Pred. No. 9e-42;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 250 LPTDCHHOCAGCTGPGKSPDCLAHNHSIGICLHPALVYNTDFEEM 301
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 232 LPTDCHHOCAGCTGPGKSPDCLAHNHSIGICLHPALVYNTDFEEM 283
 RESULT 5
 PPAP_HUMAN STANDARD; PRT; 386 AA.
 ID PPAP_HUMAN
 AC P15309;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Prostatic acid phosphatase precursor (EC 3.1.3.2).
 GN ACPp.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92272747; PubMed-1375464;
 RA Sharief F.S., Li S.S.-L.;
 RT "Structure of human prostatic acid phosphatase gene.";
 RL Biochem. Biophys. Res. Commun. 184:1468-1476(1992).
 RN [2]
 RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, DISULFIDE BONDS, AND ACTIVE
 RP SITE.
 RX MEDLINE-91115848; PubMed-1989985;
 RA van Etten R.L., Davidson R., Stevis P.E., MacArthur H., Moore D.L.;
 RT "Covalent structure, disulfide bonding, and identification of
 RT reactive surface and active site residues of human prostatic acid
 RT phosphatase.";
 RL J. Biol. Chem. 266:2313-2319(1991).
 RN [3]
 RP SEQUENCE FROM N.A.

RA MEDLINE=689228054; PubMed=2712834.
RX Sharief F.S., Lee H., Leuderman M.M., Lundvall A., Deaven L.L.,
RA Lee C.-L., Li S.S.-L.;
RT "Human prostatic acid phosphatase: cDNA cloning, gene mapping and
RT protein sequence homology with lysosomal acid phosphatase.";
RL Biochem. Biophys. Res. Commun. 160:79-86(1989).
RN [4]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Prostate;
RX MEDLINE=88312981; PubMed=2842184;
RA Vitko P., Viikkuinen P., Henttu P., Roiko K., Solin T., Huhtala M.L.;
RT "Molecular cloning and sequence analysis of cDNA encoding human
RT prostatic acid phosphatase.";
RL FEBS Lett. 236:275-281(1988).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=90370491; PubMed=2395559;
RA Tallor P.G., Govindan M.V., Patel P.C.;
RT "Nucleotide sequence of human prostatic acid phosphatase determined
RT from a full-length cDNA clone.";
RL Nucleic Acids Res. 18:4928-4928(1990).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=9508536; PubMed=7951074;
RA Sharief F.S., Li S.S.-L.;
RT "Nucleotide sequence of human prostatic acid phosphatase ACPp gene,
RT including seven Alu repeats.";
RL Biochem. Mol. Biol. Int. 33:561-565(1994).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=99023966; PubMed=9804805;
RA Lacont M.W., Handy G., Lepida L.;
RT "Structural origins of L(+)-tartarate inhibition of human prostatic
RT acid phosphatase.";
RL J. Biol. Chem. 273:30406-30409(1998).
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O - an
CC alcohol + phosphate.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SIMILARITY: BELONGS TO THE HISTIDINE ACID PHOSPHATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/sib.ch>).
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M97589; AAA60021.1; -;
DR EMBL; M97580; AAA60021.1; JOINED.
DR EMBL; M97581; AAA60021.1; JOINED.
DR EMBL; M97582; AAA60021.1; JOINED.
DR EMBL; M97583; AAA60021.1; JOINED.
DR EMBL; M97584; AAA60021.1; JOINED.
DR EMBL; M97585; AAA60021.1; JOINED.
DR EMBL; M97586; AAA60021.1; JOINED.
DR EMBL; M97587; AAA60021.1; JOINED.
DR EMBL; M97588; AAA60021.1; JOINED.
DR EMBL; M34840; AAA69694.1; -;
DR EMBL; M24902; AAA60022.1; -;
DR EMBL; X52174; CAA36422.1; -;
DR EMBL; X53605; CAA37673.1; -;
DR EMBL; U07097; AAB60640.1; -;
DR EMBL; U07083; AAB60640.1; JOINED.
DR EMBL; U07085; AAB60640.1; JOINED.
DR EMBL; U07086; AAB60640.1; JOINED.
DR EMBL; U07088; AAB60640.1; JOINED.
DR EMBL; U07091; AAB60640.1; JOINED.
DR EMBL; U07092; AAB60640.1; JOINED.
DR EMBL; U07093; AAB60640.1; JOINED.
DR EMBL; U07095; AAB60640.1; JOINED.
DR PIR; A32419; A32419; JOINED.

DR PIR: S01331; S01331.
DR PIR: S11147; S11147.
DR PIR: JH0610; JH0610.
DR PDB: ZHPA_16-SEP-98.
DR Genew; HGNC:125; ACP.
DR MIM: 171790;
DR InterPro: IPR000560; HisAc-phosphatase.
DR Pfam: PF00328; acid_phosphat_1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; 1.
KW Hydrolyase; Glycoprotein; Signal; 3d-structure.

FT SIGNAL 1
FT CHAIN 33 386
FT DISULFD 161 372
FT DISULFD 215 313
FT DISULFD 347 351
FT ACT_SITE 44 44
FT ACT_SITE 86 86
FT ACT_SITE 289 289
FT CARBOHYD 94 94
FT CARBOHYD 220 220
FT CARBOHYD 333 333
FT CONFLICT 15 24

NUCLEOPHILIC ACCEPTOR (BY SIMILARITY).
BY SIMILARITY.
PROTON DONOR.
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
N-LINKED (GLCNAC. . .).
SLGFLEFLFF -> ALASCFCFC (IN REF. 3 AND 4).
SLGFLEFLFF -> AFASCFEFC (IN REF. 5).
D -> H (IN REF. 5).
GFGQITOL -> WIPTHPA (IN REF. 4).
GEGTOLTOL -> RIMPTRHA (IN REF. 4).

E -> D (IN REF. 3).
A -> R (IN REF. 3).
O -> E (IN REF. 3).
P -> R (IN REF. 5).
P -> A (IN REF. 4).
C -> S (IN REF. 3).
S -> T (IN REF. 3).
C -> V (IN REF. 3).
D -> N (IN REF. 5).

SO SEQUENCE 386 AA; 44566 MW; EF81ELIDFAECADCA CRC64;

Query Match 5.1%; Score 35; DB 1; Length 386;
Best Local Similarity 100.0%; Pred. No. 8,9e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 MRAAPLLARAASLSLGFLLFPLFWDSDYLAKEL 35
DB 1 MRAAPLLARAASLSLGFLLFPLFWDSDYLAKEL 35

RESULT 6
CSF2_SHEEP STANDARD: PRT: 144 AA.
AC P28773;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF) (Colony-stimulating factor) (CSF).
DE CSF2.
GN Ovis aries (Sheep).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Caprinae; Ovis.
OX NCBI_Taxid=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92039044; PubMed=1937025;
RA McInnes C.J., Hald M.C.K.;
RT Cloning and expression of a cDNA encoding ovine granulocyte-macrophage colony-stimulating factor.";
RL Gene 105:1275-1279(1991).
CC -! FUNCTION: CITOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING GRANULOCYTES, MACROPHAGES, EOSINOPIHS AND ERYTHROCYTES (BY

CC SIMILARITY).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>

CC or send an email to license@sdb.ch).

CC -----

CC EMBL: X53561; CAA37632.1; -

CC PIR: JH0469; JH0469.

CC HSSP: P04141; 2GMF.

CC InterPro: IPR000773; GM_CSF.

CC Pfam: PF01109; GM_CSF; 1.

CC PRINTS: PR00693; GMC_SFCTOR.

CC PRODOM: PD007349; GM_CSF; 1.

CC SMART: SM00040; CSF2; 1.

CC PROSITE: PS00702; GM_CSF; 1.

CC Cytokine; Growth factor; Glycoprotein; Signal.

CC SIGNAL

CC CHAIN 1 17

CC BY SIMILARITY.

CC GRANULOCYTE-MACROPHAGE COLONY-STIMULATING

CC FACTOR.

CC DISULFID 71 113 BY SIMILARITY.

CC FT DISULFID 105 138 BY SIMILARITY.

CC FT CARBOHD 44 44 N-LINKED (GICNMC. . .) (POTENTIAL).

CC SEQUENCE 144 AA; 16318 MW; ABAC873B580008 CRC64;

CC

CC Query Match 3.2%; Score 22; DB 1; Length 144;

CC Best Local Similarity 100.0%; Pred. No. 1.4e-13;

CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

QY 598 DEPTCLQTRLELYKQGLRSLT 619

DB 67 DEPTCLQTRLELYKQGLRSLT 88

CC

CC RESULT 7

CC ERB2_MOUSE STANDARD: PRT; 245 AA.

CC ID ERB2_MOUSE STANDARD: PRT; 245 AA.

CC AC P70424; Q61525;

CC DT 15-DEC-1998 (Rel. 37, Created)

CC DT 15-DEC-1998 (Rel. 37, Last sequence update)

CC DT 15-JUN-2002 (Rel. 41, Last annotation update)

CC DE Receptor protein-tyrosine kinase erbB-2 (EC 2.7.1.112) (p185erbB2)

CC DE (NEU Proto-oncogene) (C-erbB-2) (Fragments).

CC GN ERB2 OR NEU.

CC OS Mus musculus (Mouse).

CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

CC OX NCBI_Taxid=10090;

CC RN [1]

CC RP SEQUENCE OF 1-149 FROM N.A.

CC RC STRAIN-CD-1; TISSUE-Uterus;

CC RX MEDLINE-97200814; PubMed-9048643;

CC RA Lin J., Day S.K., Das S.K.;

CC RT "Differential expression of the erbB2 gene in the perimplantation

CC RT mouse uterus: potential mediator of signaling by epidermal growth

CC RT factor-like growth factors.";

CC RL Endocrinology 138:1328-1337(1997).

CC [2]

CC RP SEQUENCE OF 150-245 FROM N.A.

CC RX MEDLINE-96069911; PubMed-7589796;

CC RA Moscoso L.M., Chu G.C., Gautam M., Noakes P.G., Merlie J.P.,

CC RA Sanes J.R.;

CC RT "Synapse-associated expression of an acetylcholine receptor-inducing

CC RT protein, ARIA/heretulin, and its putative receptors, ErbB2 and ErbB3,

CC RT in developing mammalian muscle.";

CC RL Dev. Biol. 172:158-169(1995).

CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,

CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A

CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-

CC ALPHA AND AMPHIREGULIN.

CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein

CC tyrosine phosphate.

CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS

CC (POTENTIAL).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN UTERINE EPITHELIAL

CC CELLS. IN THE MUSCLE, EXPRESSION LOCALIZES TO THE SYNAPTIC SITES

CC OF MUSCLE FIBERS.

CC -1- DEVELOPMENTAL STAGE: ON DAYS 1-4 OF PREGNANCY, ERB2 IS DETECTED

CC PRIMARILY IN EPITHELIAL CELLS. THE DAY 1 UTERUS SHOWING THE

CC HIGHEST ACCUMULATION. ON DAY 5, THE EPITHELIUM AND THE

CC DECIDUALIZING STROMAL CELLS AROUND THE IMPLANTING BLASTOCYST

CC EXHIBIT ACCUMULATION OF THIS RECEPTOR. ON DAYS 6-8, THE EXPRESSION

CC PERSISTS IN THE EPITHELIUM AT BOTH THE IMPLANTATION AND

CC INTERIMPLANTATION SITES IN ADDITION TO MODEST LEVELS IN THE

CC SECONDARY DECIDUAL ZONE. ON DAYS 7 AND 8, ACCUMULATION IS ALSO

CC PROMINENT IN THE TROPHOBLASTIC GIANT CELLS.

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE

CC RESIDUES (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>

CC or send an email to license@sdb.ch).

CC -----

CC EMBL: U71126; AAB17380.1; -

CC DR EMBL: L47239; AAB93532.1; -

CC DR HSSP: P11362; 1PGR.

CC MGD: MGI:95410; Erb2.

CC InterPro: IPR000719; Euk_pkinase.

CC InterPro: IPR004040; STY_pkinase.

CC InterPro: IPR001245; TYR_pkinase.

CC PRINTS: PR00109; TYRKINASE.

CC PRODOM: PD000001; Euk_kinase; 1.

CC SMART: SM00221; STYK; 1.

CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP; PARTIAL.

CC DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

CC DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.

CC KW Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;

CC KW Tyrosine-protein kinase; ATP-binding; phosphorylation.

CC FT NON_TER 1 1

CC FT DOMAIN 1 1

CC FT ACT_SITE 61 61 PROTEIN KINASE.

CC FT NON_CONS 149 150 BY SIMILARITY.

CC FT NON_TER 245 245

CC SEQUENCE 245 AA; 26927 MW; 0F763F0363DFEFC CRC64;

CC

CC Query Match 2.9%; Score 20; DB 1; Length 245;

CC Best Local Similarity 100.0%; Pred. No. 1.9e-11;

CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

QY 363 PRSPAPSGAGSDVFDGL 382

DB 210 PRSPAPSGAGSDVFDGL 229

CC

CC RESULT 8

CC CSF2_CAVPO STANDARD: PRT; 140 AA.

CC ID CSF2_CAVPO STANDARD: PRT; 140 AA.

CC AC Q60481;

CC DT 01-NOV-1997 (Rel. 35, Created)

CC DT 01-NOV-1997 (Rel. 35, Last sequence update)

CC DT 01-NOV-1997 (Rel. 35, Last annotation update)

CC DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)

CC DE (Colony-stimulating factor) (CSF) (Fragment).

CC GN CSF2.

CC DE Cavia porcellus (Guinea pig).

CC OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

CC EMBL: U22385; AAA6075.1; -
DR PIR: J10037; F0B0GM.
DR HSSP: P04141; 2GMF.
DR InterPro: IPR000773; GM_CSF.
DR Pfam: PF01109; GM_CSF; 1.
DR PRINTS: PR00693; GMCSFACR.
DR PRODOM: PD007349; GM_CSF; 1.
DR SMART: SM00040; CSF2; 1.
DR PROSITE: PS00702; GM_CSF; 1.
KM Cytokine; Growth factor; glycoprotein; signal.
FT SIGNAL 1 17
FT CHAIN 18 143
FT FT GRANULOCYTE-MACROPHAGE COLONY-STIMULATING
FT FT FACTOR.
FT DISULFID 70 112
FT FT BY SIMILARITY.
FT FT 104 137
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 143 AA; 16157 MW; 4A24E26A870A51EC CRC64;

```

Query Match 1.48; Score 10; DB 1; Length 143;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 OEPTCLOTRL 607
 DB 66 OEPTCLOTRL 75

```

RESULT 13
ZIC2_MOUSE STANDARD; PRT; 530 AA.
AC 062320;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Zinc finger protein ZIC2 (zinc finger protein of the cerebellum 2).
GN ZIC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=96132843; PubMed=8557628;
RA Aruga J., Nagai T., Tokuyama T., Hayashizaki Y., Okazaki Y.,
RA Chapman V.M., Mikoshiba K.;
RT "The mouse zic gene family. Homologues of the Drosophila pair-rule
RT gene odd-paired.";
RL J. Biol. Chem. 271:1043-1047(1996).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: CNS. A HIGH LEVEL EXPRESSION IS SEEN IN THE
CC CEREBELLUM.
CC -1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D70848; BAAL1115.1; -
CC HSSP: P08047; ISP2.
CC TRANSFAC: T04670; -
CC MCD; MGI:106679; ZIC2.
CC InterPro: IPR000822; ZnF_C2H2.
CC Pfam: PF00096; zf-C2H2; 5.
CC PRINTS: PR00046; ZINCINGER.
CC PRODOM: PD000003; ZnF_C2H2; 1.

```

```

DR SMART: SM00355; ZnF_C2H2; 4.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 4.
KM Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.
FT DOMAIN 20 23
FT DOMAIN 25 33 POLY-ALA.
FT DOMAIN 89 97 POLY-ALA.
FT DOMAIN 227 231 POLY-ALA.
FT DOMAIN 232 239 POLY-HIS.
FT DOMAIN 300 415 ZINC_FINGERS.
FT ZN_FING 300 327 C2H2-TYPE (ATYPICAL).
FT ZN_FING 333 357 C2H2-TYPE.
FT ZN_FING 363 387 C2H2-TYPE.
FT ZN_FING 393 415 C2H2-TYPE.
FT ZN_FING 456 470 POLY-ALA.
FT DOMAIN 501 512 POLY-ALA.
SQ SEQUENCE 530 AA; 55492 MW; 0065BD75B52E7DD2 CRC64;

```

Query Match 1.38; Score 9; DB 1; Length 530;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 682 AAAHHHHH 690
 DB 229 AAAHHHHH 237

```

RESULT 14
ZIC2_HUMAN STANDARD; PRT; 532 AA.
AC 095409; O9H309.
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zinc finger protein ZIC2 (zinc finger protein of the cerebellum 2).
GN ZIC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND VARIANT HPES POLY-ALA INSERTION.
RX MEDLINE=98442655; PubMed=10984499;
RA Brown S.A., Warburton D., Brown L.Y., Yu C.Y., Roeder E.R.,
RA Stengel-Rutkowski S., Hennekam R.C., Muenke M.;
RT "Holoencephaly due to mutations in ZIC2, a homologue of Drosophila
RT odd-paired.";
RL Nat. Genet. 20:180-183(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20556339; PubMed=10984499;
RA Yang Y., Hwang C.K., Junn E., Lee G., Mouradian M.M.;
RT "ZIC2 and Sp3 repress Sp1-induced activation of the human D1A dopamine
RT receptor gene.";
RL J. Biol. Chem. 275:38863-38869(2000).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DISEASE: DEFECTS IN ZIC2 ARE A CAUSE OF HOLOPROSENCEPHALY TYPE 5
CC (HPES). HPES IS A STRUCTURAL ANOMALY OF THE BRAIN.
CC -1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF104902; AAC96325.1; -
CC DR EMBL: AF193855; AAC28409.1; -
CC HSSP: P08047; ISP2.

```


THIS PAGE BLANK (USPTO)


```

QY 98 EVQGYVLIANOVROVPLQRLIRVGTQLEFEDNVALAVLNDGDLNNTTPTVGASPGGLR 157
    |||||||
DB 79 EVQGYVLIANOVROVPLQRLIRVGTQLEFEDNVALAVLNDGDLNNTTPTVGASPGGLR 138
QY 158 ELQRLSTLEIKGVLQIQRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPM 217
    |||||||
DB 139 ELQRLSTLEIKGVLQIQRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPM 198
QY 218 CKGRSGWSESDCQSLTRVYVCGGCAKCGKPLPTDCCHOCAGCGPHSDCLAHF 277
    |||||||
DB 199 CKGRSGWSESDCQSLTRVYVCGGCAKCGKPLPTDCCHOCAGCGPHSDCLAHF 258
QY 278 NMSGICELHCPALVTYNTDFTFESMPNDEGRYTFEGASCVTACPYNYLSTDVGS 329
    |||||||
DB 259 NMSGICELHCPALVTYNTDFTFESMPNDEGRYTFEGASCVTACPYNYLSTDVGS 310

RESULT 2
Q14256 PRELIMINARY: PRT: 165 AA.
ID Q14256:
AC Q14256:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE C-erb B2/neu protein (Fragment).
GN C-ERB B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=6070181; Pubmed=2999974;
RA Consens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
RA Seeburg P.H., Liberman T.A., Schlessinger J., Franke U.,
RA Levanon A., Ullrich A.;
RT Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.;
RL Science 230:1132-1139(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=94000386; Pubmed=8104414;
RA Sarkar F.H., Ball D.E., Li Y.W., Cissman J.D.;
RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERB2)
RT gene.";
RL DNA Cell Biol. 12:611-615(1993).
DR EMBL: M95667; AAC37531.1; -.
FT NON_TER 1
SQ SEQUENCE 165 AA; 17327 MW; A0C113BA308BF46B CRC64;

Query Match 23.9%; Score 165; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.2e-159;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 373 AGSDVFEFGDGLGMAKAGLSLPHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 432
    |||||||
DB 1 AGSDVFEFGDGLGMAKAGLSLPHDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV 60
QY 433 NOPDVRRQPPSPREGPLPAARPAAGATLERAKTISPGKNGVYKQVAFAGCAVENPEYLTQP 492
    |||||||
DB 61 NOPDVRRQPPSPREGPLPAARPAAGATLERAKTISPGKNGVYKQVAFAGCAVENPEYLTQP 120
QY 493 GGAAPQHPHPAPAFADNLTYWDQDPERGAPSTFKGTPTAEN 537
    |||||||
DB 121 GGAAPQHPHPAPAFADNLTYWDQDPERGAPSTFKGTPTAEN 165

RESULT 3
Q8WYV0 PRELIMINARY: PRT: 412 AA.
ID Q8WYV0:
AC Q8WYV0:
DT 01-MAR-2002 (TReMBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 44.7 kDa protein.
GN p3659.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.O., Qin W.X., Zhao X.T.,
RA Wan D.F., Gu J.R.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF318349; AAL5856.1; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR00719; Euk_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PFO0069; VLP_motif.
DR Pfam: PFO2757; VLP_2.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00219; TYKc; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Hypothetical protein.
SO SEQUENCE 412 AA; 44702 MW; 034397FF3F27D2BC CRC64;

Query Match 19.18%; Score 132; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.2e-125;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 GAGCMVHHRRSSSTRSGGDLTLGLPSEEARSPPLAPSEGASDVFQDGLGMAK 389
    |||||||
DB 150 GAGCMVHHRRSSSTRSGGDLTLGLPSEEARSPPLAPSEGASDVFQDGLGMAK 209
QY 390 LQSLPTHDPSPLOKRYSDPTVPLPSETDGYVAPLTCSPQPEYVQPDVRRQPPSPREGPL 449
    |||||||
DB 210 LQSLPTHDPSPLOKRYSDPTVPLPSETDGYVAPLTCSPQPEYVQPDVRRQPPSPREGPL 269
QY 450 PAARPAAGATLER 461
    |||||||
DB 270 PAARPAAGATLER 281

RESULT 4
Q9BG66 PRELIMINARY: PRT: 149 AA.
ID Q9BG66:
AC Q9BG66:
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Receptor tyrosine kinase ErbB2 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
RA Tectens F., Fischer B.;
RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
RT peri-implantation rabbit uterus and blastocyst.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF333178; AAK14371.1; -.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PFO0757; Furin-like; 1.
DR SMART: SM00261; FU; 2.
DR Kinase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;

```

Query Match 9.7%; Score 67; DB 6; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.3e-59;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 250 LPTDCHEQCAAGCTGPRKSHDCLACLFHNSGICELHCPALVYNTDTFESMPNEGRTT 309
 DB 74 LPTDCHEQCAAGCTGPRKSHDCLACLFHNSGICELHCPALVYNTDTFESMPNEGRTT 133

OY 310 FGASCVT 316
 DB 134 FGASCVT 140

RESULT 5
 ID 018735 PRELIMINARY; PRT; 1259 AA.
 AC 018735;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE EtRB-2;
 OS Carls familiaris (dog)
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yokota H.;
 RT "cDNA cloning of etrb-2 from canine mammary gland."
 RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB008451; BAA23127.1;
 DR HSSP; P11362; IFGK.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000494; EGFR_L_domain.
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR002174; Furin-like.
 DR InterPro; IPR001245; Tyr_Pkinase.
 DR InterPro; IPR004019; YLP_motif.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 DR Pfam; PF02757; YLP; 2.
 DR ProDom; PD000001; Euk_Pkinase; 1.
 DR SMART; SM00261; FU; 3.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KM ATP-binding; Transferase; Tyrosine-protein kinase.
 SQ SEQUENCE 1259 AA; 137968 MW; E37364D49C4ACD46 CRC64;

Query Match 9.6%; Score 66; DB 6; Length 1259;
 Best Local Similarity 100.0%; Pred. No. 1.1e-57;
 Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 PTDCHEQCAAGCTGPRKSHDCLACLFHNSGICELHCPALVYNTDTFESMPNEGRTT 310
 DB 232 PTDCHEQCAAGCTGPRKSHDCLACLFHNSGICELHCPALVYNTDTFESMPNEGRTT 291

OY 311 GASCVT 316
 DB 292 GASCVT 297

RESULT 6
 ID 09GL44 PRELIMINARY; PRT; 144 AA.
 AC 09GL44;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Granulocyte-macrophage colony-stimulating factor.
 GN GM-CSF.
 OS Macaca mulatta (Rhesus macaque).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hutchinson K.L., Villinger F., Miranda M.E., Ksiazek T.G.,
 RA Peters C.J., Rollin P.E.;
 RT "Multiplex analysis of cytokines in the sera of cynomolgus macaques
 RT naturally infected with Ebola (Reston)."
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY007376; AAC16526.1;
 DR HSSP; P04141; 2GMF.
 DR InterPro; IPR00773; GM_CSF.
 DR Pfam; PF01109; GM_CSF; 1.
 DR PRINTS; PR00693; GMCSFACTOR.
 DR ProDom; PD007349; GM_CSF; 1.
 DR SMART; SM00040; CSF2; 1.
 DR PROSITE; PS00702; GM_CSF; 1.
 FT VARIANT 60 60 V -> I.
 SQ SEQUENCE 144 AA; 16177 MW; 7D5F381DA2FC832P CRC64;

Query Match 5.2%; Score 36; DB 6; Length 144;
 Best Local Similarity 100.0%; Pred. No. 4.8e-28;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 616 GSLTKKGPRTWMASHYKHCPRPTSCAQITTF 651
 DB 85 GSLTKKGPRTWMASHYKHCPRPTSCAQITTF 120

RESULT 7
 ID 096KY0 PRELIMINARY; PRT; 386 AA.
 AC 096KY0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Acid phosphatase, prostatic.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue-Prostate;
 RC Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; BC016344; AAH16344.1;
 DR InterPro; IPR000560; HsAc_phsphtse.
 DR Pfam; PF00328; acid_phosphat; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 SQ SEQUENCE 386 AA; 44540 MW; FE90E10CBCEACDA CRC64;

Query Match 5.1%; Score 35; DB 4; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.3e-26;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRAAPLLARAASISLGFLLFPWLDRLAKEL 35
 DB 1 MRAAPLLARAASISLGFLLFPWLDRLAKEL 35

RESULT 8
 ID 096QMO PRELIMINARY; PRT; 418 AA.
 AC 096QMO;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Acid phosphatase, prostate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Strauberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC007460; AAH07460.1; -
DR InterPro: IPR000560; HisAc.phosphatase.
DR Pfam: PF00328; acid_phosphatase.1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN.1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN.1.
SQ SEQUENCE 418 AA; 48308 MW; 168E10406974E4462 CRC64;

Query Match 5.1%; Score 35; DB 4; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.4e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRAAPLLARAASLSLGFLLFFLFDLSVLAKEI 35
DB 1 MRAAPLLARAASLSLGFLLFFLFDLSVLAKEI 35

RESULT 9

AC 08R2X1 PRELIMINARY; PRT; 367 AA.
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Hypothetical 40.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC027080; AAH27080.1; -
KW Hypothetical protein.
SQ SEQUENCE 367 AA; 40163 MW; 0BE03395F9E101B0 CRC64;

Query Match 3.9%; Score 27; DB 11; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.7e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 463 KTLSPGKNGVYKDYFAFGAVENPEYL 489
DB 283 KTLSPGKNGVYKDYFAFGAVENPEYL 309

RESULT 10

ID 09MYK4 PRELIMINARY; PRT; 144 AA.
AC 09MYK4;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Granulocyte-macrophage colony-stimulating factor (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
ON NCBI_TaxID=9940;
RX [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91331592; PubMed=1869289;
O'Brien P.M., Rothel J.S., Seow H.F., Wood P.R.;

RT "Cloning and sequencing of the cDNA for ovine granulocyte-macrophage
 colony stimulating factor (GM-CSF).";
RL Immunol. Cell Biol. 69:51-55(1991).
DR EMBL: X55991; CA39463.1; -
DR HSSP: P04141; 2GMF
DR InterPro: IPR000773; GM-CSF.
DR Pfam: PF01109; GM-CSF.1.
DR PRINTS: PR00693; GMCSFACOR.
DR PRODOM: PD007349; GM-CSF.1.
DR SMART: SM00040; CSF2.1.
DR PROSITE: PS00702; GM-CSF.1.
FT NON_TER 144 144
SQ SEQUENCE 144 AA; 16290 MW; ABAAD7633B580008 CRC64;

Query Match 3.2%; Score 22; DB 6; Length 144;
Best Local Similarity 100.0%; Pred. No. 8.6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 QEPICLOTRLELYKQGLRSLT 619
DB 67 QEPICLOTRLELYKQGLRSLT 88

RESULT 11

ID 096OK9 PRELIMINARY; PRT; 386 AA.
AC 096OK9;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Acid phosphatase, prostate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Strauberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC008493; AAH08493.1; -
DR InterPro: IPR000560; HisAc.phosphatase.
DR Pfam: PF00328; acid_phosphatase.1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN.1.
DR PROSITE: PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN.1.
SQ SEQUENCE 386 AA; 44515 MW; AAD817CEC1DC1A84 CRC64;

Query Match 2.9%; Score 20; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2.4e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LGFLFLFFWLDRLSLAKEL 35
DB 16 LGFLFLFFWLDRLSLAKEL 35

RESULT 12

ID 08WN17 PRELIMINARY; PRT; 146 AA.
AC 08WN17;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Granulocyte-macrophage colony-stimulating factor (Fragment).
ON GM-CSF.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
ON NCBI_TaxID=9796;
RX [1]
RP SEQUENCE FROM N.A.
RA Vecchione A., D'Amelio F., Kanellos T.S., Howard C.R., Hamblin A.S.,
Catchpole B.;

RT "cDNA for equine GM-CSF."
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF448481; AAL1017.1;
 DR InterPro: IPR000773; GM-CSF.
 DR Pfam: PF01109; GM-CSF; 1.
 DR PRINTS: PR00693; GMCSFACTOR.
 DR ProDom: PD007349; GM-CSF; 1.
 DR SMART: SM00040; CSF2; 1.
 DR PROSITE: PS00702; GM-CSF; UNKNOWN_1.
 FT NON-TER 146 146
 SQ SEQUENCE 146 AA: 16594 MW: 162C19152391281E CRC64;

Query Match 2.5%; Score 17; DB 6; Length 146;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 623 GPLTMASHYKHCPT 639
 DB 92 GPLTMASHYKHCPT 108

RESULT 13

O95L10 PRELIMINARY: PRT; 152 AA.
 AC O95L10;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Granulocyte-macrophage colony-stimulating-factor.
 GN GM-CSF.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Manuel S., Commandeur U., Steinhach F.;
 RT "Cloning of equine granulocyte-macrophage colony-stimulating-factor
 (eq-GM-CSF)."
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY040203; AAK72108.2;
 DR InterPro: IPR000773; GM-CSF.
 DR Pfam: PF01109; GM-CSF; 1.
 DR ProDom: PD007349; GM-CSF; 1.
 DR PROSITE: PS00702; GM-CSF; UNKNOWN_1.
 SQ SEQUENCE 152 AA: 17173 MW: 75605CC1ADE9FE9 CRC64;

Query Match 2.5%; Score 17; DB 6; Length 152;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 623 GPLTMASHYKHCPT 639
 DB 92 GPLTMASHYKHCPT 108

RESULT 14

O99J91 PRELIMINARY: PRT; 138 AA.
 AC O99J91;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Granulocyte-macrophage colony stimulating factor precursor.
 OS Marmota monax (Woodchuck).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
 OC Marmota.
 OX NCBI_TaxID=9995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wu H.-L., Chen P.-J., Lin H.-K., Lee R.-S., Lin H.-L., Chen D.-S.;
 RT "Molecular Cloning and Expression of Woodchuck Granulocyte-Macrophage

RT Colony Stimulating Factor."
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF255734; AAG49541.1;
 DR EMBL: AF255735; AAG49542.1;
 DR HSRF: P04141; 2GMF.
 DR InterPro: IPR000773; GM-CSF.
 DR Pfam: PF01109; GM-CSF; 1.
 DR PRINTS: PR00693; GMCSFACTOR.
 DR ProDom: PD007349; GM-CSF; 1.
 DR SMART: SM00040; CSF2; 1.
 DR PROSITE: PS00702; GM-CSF; 1.
 FT SIGNAL.
 FT SIGNAL.
 FT CHAIN 20 138
 FT CHAIN
 SQ SEQUENCE 138 AA: 15297 MW: A6B9E04035A806FD CRC64;

Query Match 1.3%; Score 9; DB 11; Length 138;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 599 EPTCLOTRL 607
 DB 68 EPTCLOTRL 76

RESULT 15

O8XU41 PRELIMINARY: PRT; 253 AA.
 AC O8XU41;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical transmembrane protein RSC3353.
 GN RSC3353 OR R502628.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM11000;
 RX MEDLINE-21681879; PubMed-11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Cartolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigler P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
 RL Nature 415:497-502(2002).
 DR EMBL: AL646074; CAD17141.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 253 AA: 26668 MW: 95DBE888F9F0B58 CRC64;

Query Match 1.3%; Score 9; DB 16; Length 253;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 LILARRASL 14
 DB 68 LILARRASL 76

Search completed: April 28, 2003, 13:45:48
 Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:44:15 ; Search time 17 Seconds
(without alignments)
1194.224 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690

Sequence: 1 MRAAPLLARAASLSLGLF.....EPVGGAPRRRAAHNNHHN 690

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents-AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTC05.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	292	42.3	419	4	US-09-630-155-2 Sequence 2, Appl
2	292	42.3	782	2	US-09-146-283-4 Sequence 4, Appl
3	292	42.3	782	3	US-08-579-823A-4 Sequence 4, Appl
4	292	42.3	782	1	US-09-344-195-4 Sequence 4, Appl
5	292	42.3	1255	1	US-08-467-083-68 Sequence 68, Appl
6	292	42.3	1255	1	US-08-414-417B-68 Sequence 68, Appl
7	292	42.3	1255	2	US-08-484-438-8 Sequence 8, Appl
8	292	42.3	1255	2	US-08-486-348A-68 Sequence 8, Appl
9	292	42.3	1255	2	US-08-625-101-2 Sequence 2, Appl
10	292	42.3	1255	2	US-08-468-545B-68 Sequence 68, Appl
11	292	42.3	1255	2	US-08-356-786-2 Sequence 2, Appl
12	292	42.3	1255	3	US-08-466-680B-68 Sequence 68, Appl
13	289	41.9	624	3	US-08-422-108-1 Sequence 1, Appl
14	289	41.9	624	4	US-08-422-734-1 Sequence 1, Appl
15	132	19.1	580	1	US-08-414-417B-69 Sequence 69, Appl
16	132	19.1	580	2	US-08-486-348A-69 Sequence 69, Appl
17	132	19.1	580	2	US-08-468-545B-69 Sequence 69, Appl
18	132	19.1	580	3	US-08-466-680B-69 Sequence 69, Appl
19	128	18.6	165	1	US-08-318-193-8 Sequence 8, Appl
20	128	18.6	165	1	US-08-318-193-10 Sequence 10, Appl
21	127	18.4	127	1	US-08-318-193-2 Sequence 2, Appl
22	127	18.4	127	6	5229496-15 Patent No. 5229496
23	127	18.4	128	3	US-08-469-318-160 Sequence 160, App
24	127	18.4	128	3	US-08-468-609A-160 Sequence 160, App
25	127	18.4	128	4	US-08-446-872A-160 Sequence 160, App
26	127	18.4	128	4	US-08-762-227A-160 Sequence 160, App
27	127	18.4	128	5	PCT-US95-01185-160 Sequence 160, App

28	127	18.4	131	6	5229496-2 Patent No. 5229496
29	127	18.4	144	1	US-08-284-393B-11 Sequence 11, Appl
30	127	18.4	144	4	US-09-522-217-114 Sequence 11, Appl
31	127	18.4	144	5	PCT-US95-08950-11 Sequence 11, Appl
32	127	18.4	259	3	US-08-469-318-141 Sequence 141, App
33	127	18.4	259	3	US-08-468-609A-141 Sequence 141, App
34	127	18.4	259	4	US-08-446-872A-141 Sequence 141, App
35	127	18.4	259	4	US-08-762-227A-141 Sequence 141, App
36	127	18.4	259	5	PCT-US95-01185-141 Sequence 141, App
37	127	18.4	274	3	US-08-469-318-144 Sequence 144, App
38	127	18.4	274	3	US-08-468-609A-144 Sequence 144, App
39	127	18.4	274	4	US-08-446-872A-144 Sequence 144, App
40	127	18.4	274	4	US-08-762-227A-144 Sequence 144, App
41	127	18.4	301	3	PCT-US95-01185-144 Sequence 144, App
42	127	18.4	301	3	US-08-469-318-142 Sequence 142, App
43	127	18.4	301	3	US-08-468-609A-142 Sequence 142, App
44	127	18.4	301	4	US-08-446-872A-142 Sequence 142, App
45	127	18.4	301	4	US-08-762-227A-142 Sequence 142, App

ALIGNMENTS

RESULT 1
US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130
GENERAL INFORMATION:
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: DAVIS WRIGHT TREMINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2
Query Match 42.3%; Score 292; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 2e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 GASTGVCTGTDKLLKLPSPETHLDMRLRLXGCGVGVGNLELTLPNASTSPFDIO 97
DB 19 GAATGVCTGTDKLLKLPSPETHLDMRLRLXGCGVGVGNLELTLPNASTSPFDIO 78
QY 98 EVGVYLLIANQVRQVPLRLRIVRCTOLFEDNYALAVLDNGDPLNNTTPVVGASPGGLR 157

Db 79 EVGGVILAHNOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138
Qy 158 ELQLSRLEILKGVLIQIORNPOLCYODITLTKMDFHKNQNALTLIDTNRSRACHPCSPM 217
Db 139 ELQLSRLEILKGVLIQIORNPOLCYODITLTKMDFHKNQNALTLIDTNRSRACHPCSPM 198
Qy 218 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 277
Db 199 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 258
Qy 278 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNTLSDVGS 329
Db 259 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNTLSDVGS 310

RESULT 2
US-09-146-283-4
Sequence 4, Application US/09146283
Patent No. 5976546

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146.283
FILING DATE: 03-SEPT-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0980
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4

Query Match 42.3%; Score 292; DB 2; Length 782;
Best Local Similarity 100.0%; Pred. No. 3.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GAATOVCTGTDMKRLRASPETHLDMRLHLYGCGVVOGNGLELYLPTNLSLFLDIO 97
Db 19 GAATOVCTGTDMKRLRASPETHLDMRLHLYGCGVVOGNGLELYLPTNLSLFLDIO 78
Qy 98 EVGGVILAHNOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 157
Db 79 EVGGVILAHNOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138

Qy 158 ELQLSRLEILKGVLIQIORNPOLCYODITLTKMDFHKNQNALTLIDTNRSRACHPCSPM 217
Db 139 ELQLSRLEILKGVLIQIORNPOLCYODITLTKMDFHKNQNALTLIDTNRSRACHPCSPM 198
Qy 218 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 277
Db 199 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGPLPTDCCHEGCAAGCTGPKHSDCLACLHF 258
Qy 278 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNTLSDVGS 329
Db 259 NMSGICELHCPALVTYNTDTEFSMPNPEGRTYFGASCYACPYNTLSDVGS 310

RESULT 3
US-08-579-823A-4
Sequence 4, Application US/08579823A
Patent No. 6080409

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Composition and Method
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Denlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/579.823A
FILING DATE: 03-DEC-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0980
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match 42.3%; Score 292; DB 3; Length 782;
Best Local Similarity 100.0%; Pred. No. 3.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GAATOVCTGTDMKRLRASPETHLDMRLHLYGCGVVOGNGLELYLPTNLSLFLDIO 97
Db 19 GAATOVCTGTDMKRLRASPETHLDMRLHLYGCGVVOGNGLELYLPTNLSLFLDIO 78
Qy 98 EVGGVILAHNOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 157
Db 79 EVGGVILAHNOVROVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138
Qy 158 ELQLSRLEILKGVLIQIORNPOLCYODITLTKMDFHKNQNALTLIDTNRSRACHPCSPM 217
Db 139 ELQLSRLEILKGVLIQIORNPOLCYODITLTKMDFHKNQNALTLIDTNRSRACHPCSPM 198

|||||
Db 199 CKGSRCHGESSBDCQSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 258
Qy 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 329
Db 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 310

RESULT 6
US-08-414-17B-68
Sequence 68, Application US/08414417B
Patent No. 5801005
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/414,417B
FILING DATE: 31-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-414-17B-68

Query Match 42.3%; Score 292; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GAASTGVCTGDMKRLRPASPEHLDMLRHLYGCGVYVGNLELTYLPTNASLSFLDIO 97
Db 19 GAASTGVCTGDMKRLRPASPEHLDMLRHLYGCGVYVGNLELTYLPTNASLSFLDIO 78
Qy 98 EVGGYVLIANQVROVPLQRLRIYRGTLQFEDNYALVLDNGDPLNNTPTVTGASPGGLR 157
Db 79 EVGGYVLIANQVROVPLQRLRIYRGTLQFEDNYALVLDNGDPLNNTPTVTGASPGGLR 138
Qy 158 ELQARSLTEILKGGVLIQNRNOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPM 217
Db 139 ELQARSLTEILKGGVLIQNRNOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPM 198
Qy 218 CKGSRCHGESSBDCQSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 277
Db 199 CKGSRCHGESSBDCQSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 258
Qy 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 329
Db 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 310

|||||
Db 199 CKGSRCHGESSBDCQSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 258
Qy 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 329
Db 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 310

RESULT 7
US-08-484-438-8
Sequence 8, Application US/08484438
Patent No. 5811098
Patent No. 5811098 5780031
GENERAL INFORMATION:
APPLICANT: PLOWMAN, Gregory D.
APPLICANT: CULOUSCOU, Jean-Michel
APPLICANT: SHOYAB, Mohammed
APPLICANT: SIEGALL, Clay B.
APPLICANT: HELISTR m, Ingegerd
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: MISTROCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-438-8

Query Match 42.3%; Score 292; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GAASTGVCTGDMKRLRPASPEHLDMLRHLYGCGVYVGNLELTYLPTNASLSFLDIO 97
Db 19 GAASTGVCTGDMKRLRPASPEHLDMLRHLYGCGVYVGNLELTYLPTNASLSFLDIO 78
Qy 98 EVGGYVLIANQVROVPLQRLRIYRGTLQFEDNYALVLDNGDPLNNTPTVTGASPGGLR 157
Db 79 EVGGYVLIANQVROVPLQRLRIYRGTLQFEDNYALVLDNGDPLNNTPTVTGASPGGLR 138
Qy 218 CKGSRCHGESSBDCQSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 277
Db 199 CKGSRCHGESSBDCQSLRTYCAGGCAKCGPLPTDCHECCAGCTGPKHSDCLCLHF 258
Qy 278 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 329
Db 259 NHSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNLTSDVGS 310

Oy 218 CKGSRCKWESSSEDCSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 277
Db 199 CKGSRCKWESSSEDCSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 258

Oy 278 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329
Db 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310

RESULT 8
US-08-486-348A-68
Sequence 68, Application US/08486348A
Patent No. 5846538

GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match 42.3%; Score 292; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 38 GAASOVCTGTDMLRLPASPETHLDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 97
Db 19 GAASOVCTGTDMLRLPASPETHLDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 78
Oy 98 EVOGVYVLAHNVQVPLQRLRIVRGTOLEFEDNVYLAVALDNDPLNNTTPVYGASPGGLR 157
Db 79 EVOGVYVLAHNVQVPLQRLRIVRGTOLEFEDNVYLAVALDNDPLNNTTPVYGASPGGLR 138
Oy 158 ELQLSLTEILKGVLIQVLIQNPOLCYODITLMKDIFHKNNQALTLTIDNRSRACHPCSPM 217
Db 139 ELQLSLTEILKGVLIQVLIQNPOLCYODITLMKDIFHKNNQALTLTIDNRSRACHPCSPM 198
Oy 218 CKGSRCKWESSSEDCSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 277
Db 199 CKGSRCKWESSSEDCSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 258
Oy 278 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329
Db 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310

Db 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310

RESULT 9
US-08-625-101-2
Sequence 2, Application US/08625101
Patent No. 5869445

GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2

Query Match 42.3%; Score 292; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 38 GAASOVCTGTDMLRLPASPETHLDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 97
Db 19 GAASOVCTGTDMLRLPASPETHLDMLRHLVQGCVOVGNLELYLPTNASTLSFLQDIO 78
Oy 98 EVOGVYVLAHNVQVPLQRLRIVRGTOLEFEDNVYLAVALDNDPLNNTTPVYGASPGGLR 157
Db 79 EVOGVYVLAHNVQVPLQRLRIVRGTOLEFEDNVYLAVALDNDPLNNTTPVYGASPGGLR 138
Oy 158 ELQLSLTEILKGVLIQVLIQNPOLCYODITLMKDIFHKNNQALTLTIDNRSRACHPCSPM 217
Db 139 ELQLSLTEILKGVLIQVLIQNPOLCYODITLMKDIFHKNNQALTLTIDNRSRACHPCSPM 198
Oy 218 CKGSRCKWESSSEDCSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 277
Db 199 CKGSRCKWESSSEDCSLRTVCAGGCARCKGPLPTDCCHCCAGCTGPKHSDCLACLHF 258
Oy 278 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 329
Db 259 NHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYACPYNYLSTDVGS 310
RESULT 10
US-08-468-545B-68

Sequence 68, Application US/08468545B
Patent No. 5876712
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 20010.448C5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match 42.3%; Score 292; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 38 GAATGCTGDMKLRPASPETHLDMRLHYGCGVVGNGLELYLPTNLSFLDDIQ 97
DB 19 GAATGCTGDMKLRPASPETHLDMRLHYGCGVVGNGLELYLPTNLSFLDDIQ 78
DB 98 EVGGYVLIHQVQVPLQRLRIYRGTOLEFEDNALAVLDGDDPLNNTPTVTGASPGGLR 157
DB 79 EVGGYVLIHQVQVPLQRLRIYRGTOLEFEDNALAVLDGDDPLNNTPTVTGASPGGLR 138
DB 158 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
DB 218 CKGRKCGESEDQSLRTFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACHF 277
DB 199 CKGRKCGESEDQSLRTFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACHF 258
DB 278 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 11
US-08-356-786-2
Sequence 2, Application US/08356786
Patent No. 5877305
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69

TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-356-786-2

Query Match 42.3%; Score 292; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.5e-269;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 38 GAATGCTGDMKLRPASPETHLDMRLHYGCGVVGNGLELYLPTNLSFLDDIQ 97
DB 19 GAATGCTGDMKLRPASPETHLDMRLHYGCGVVGNGLELYLPTNLSFLDDIQ 78
DB 98 EVGGYVLIHQVQVPLQRLRIYRGTOLEFEDNALAVLDGDDPLNNTPTVTGASPGGLR 157
DB 79 EVGGYVLIHQVQVPLQRLRIYRGTOLEFEDNALAVLDGDDPLNNTPTVTGASPGGLR 138
DB 158 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
DB 218 CKGRKCGESEDQSLRTFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACHF 277
DB 199 CKGRKCGESEDQSLRTFYCAGGACRCKGRLPTDCHECCAGCGPKHSDCLACHF 258
DB 278 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 12
US-08-466-680B-68
Sequence 68, Application US/08466680B
Patent No. 6075122
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69

APPLICANT: Hudziak, Robert M.
 APPLICANT: Shepherd, H. Michael
 APPLICANT: Ullrich, Axel
 TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Genentech, Inc.,
 STREET: 460 Point San Bruno Blvd
 CITY: South San Francisco
 STATE: California
 COUNTRY: USA

```

: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Winpatlin (Genetech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/422,734
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/422108
: FILING DATE: 14-Apr-1995
: APPLICATION NUMBER: 08/355460
: FILING DATE: 13-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/048346
: FILING DATE: 15-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/354319
: FILING DATE: 19-MAY-1989
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M
: REGISTRATION NUMBER: 00,000
: REFERENCE/DOCKET NUMBER: 554C2D1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 624 amino acids
: TYPE: Amino Acid
: TOPOLOGY: Linear
:
: US-08-422-734-1
:
: Query Match 41.9%; Score 289; DB 4; Length 624;
: Best Local Similarity 100.0%; Pred. No. 2e-286;
: Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 41 STQCTGTDMKRLPASPETHLDMRLHYOGCQVVGKLELYLPTNASLSFLDIDIOEQ 100
:   |||
: Db 1 STQCTGTDMKRLPASPETHLDMRLHYOGCQVVGKLELYLPTNASLSFLDIDIOEQ 60
:
: QY 101 GYVLIANOVQVPLQRLRYRGTOLEFDNYALAVLNDGDPNLTTPVTGASPGGLRELQ 160
:   |||
: Db 61 GYVLIANOVQVPLQRLRYRGTOLEFDNYALAVLNDGDPNLTTPVTGASPGGLRELQ 120
:
: QY 161 LRSLEILKGGVLIQRPOLCYODTILMKDIFHNNOLALTLIDTNSRACHPCSPCKG 220
:   |||
: Db 121 LRSLEILKGGVLIQRPOLCYODTILMKDIFHNNOLALTLIDTNSRACHPCSPCKG 180
:
: QY 221 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHFNHS 280
:   |||
: Db 181 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHFNHS 240
:
: QY 281 GICELHCPALVTYNTDIFESMPNPEGRTTGASCVTACPYNYLSTDVGS 329
:   |||
: Db 241 GICELHCPALVTYNTDIFESMPNPEGRTTGASCVTACPYNYLSTDVGS 289
:
: RESULT 15
: US-08-414-417B-69
: Sequence 69, Application US/08414417B
: Patent No. 5801005
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
: TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
: NUMBER OF SEQUENCES: 69
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESSEE: Seed and Berry LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/414,417B
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 69:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 580 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
:
: US-08-414-417B-69
:
: Query Match 19.1%; Score 132; DB 1; Length 580;
: Best Local Similarity 100.0%; Pred. No. 4.1e-117;
: Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 330 GAGGVMVHRRSSSTRSGGDLTLGLEPSEEPSPPLAPSEGASDVDFDGLGGAAG 389
:   |||
: Db 363 GAGGVMVHRRSSSTRSGGDLTLGLEPSEEPSPPLAPSEGASDVDFDGLGGAAG 422
:
: QY 390 LQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVQPDVRRQPPSPREGPL 449
:   |||
: Db 423 LQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVQPDVRRQPPSPREGPL 482
:
: QY 450 PAARPAGATLER 461
:   |||
: Db 483 PAARPAGATLER 494
:
: Search completed: Apr-11 28, 2003, 13:46:38
: Job time: 19 secs

```

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:44:40 ; Search time 22 Seconds
(without alignments)
2513.162 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690

Sequence: 1 MRAAPLLARAASLSLGLF.....EPVQCAPPPAAHHHHH 690

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 301932 seqs, 80129803 residues

Word size: 0

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published_Applications_AA:*

1: /cgn2-6/ptodata/1/pubppaa/US08_NEW_PUB.pcp:*
2: /cgn2-6/ptodata/1/pubppaa/PCT_NEW_PUB.pcp:*
3: /cgn2-6/ptodata/1/pubppaa/US06_NEW_PUB.pcp:*
4: /cgn2-6/ptodata/1/pubppaa/US07_PUBCOMB.pcp:*
5: /cgn2-6/ptodata/1/pubppaa/US07_PUBCOMB.pcp:*
6: /cgn2-6/ptodata/1/pubppaa/US07_PUBCOMB.pcp:*
7: /cgn2-6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pcp:*
8: /cgn2-6/ptodata/1/pubppaa/US08_PUBCOMB.pcp:*
9: /cgn2-6/ptodata/1/pubppaa/US09_NEW_PUB.pcp:*
10: /cgn2-6/ptodata/1/pubppaa/US09_PUBCOMB.pcp:*
11: /cgn2-6/ptodata/1/pubppaa/US10_NEW_PUB.pcp:*
12: /cgn2-6/ptodata/1/pubppaa/US10_PUBCOMB.pcp:*
13: /cgn2-6/ptodata/1/pubppaa/US60_NEW_PUB.pcp:*
14: /cgn2-6/ptodata/1/pubppaa/US60_PUBCOMB.pcp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	690	100.0	690	10 US-09-821-883-2	Sequence 2, Appl1
2	549	79.6	555	10 US-09-821-883-1	Sequence 1, Appl1
3	329	47.7	479	10 US-09-821-883-5	Sequence 5, Appl1
4	329	47.7	564	10 US-09-821-883-3	Sequence 3, Appl1
5	329	47.7	697	10 US-09-821-883-4	Sequence 4, Appl1
6	292	42.3	645	10 US-09-921-161-1	Sequence 1, Appl1
7	292	42.3	653	9 US-09-854-356-3	Sequence 3, Appl1
8	292	42.3	712	9 US-09-854-356-7	Sequence 7, Appl1
9	292	42.3	919	9 US-09-854-356-6	Sequence 6, Appl1
10	292	42.3	1255	9 US-09-769-508-2	Sequence 2, Appl1
11	292	42.3	1255	9 US-09-854-356-1	Sequence 1, Appl1
12	292	42.3	1255	9 US-09-930-125-2	Sequence 2, Appl1
13	292	42.3	1255	9 US-09-441-411-6	Sequence 6, Appl1
14	292	42.3	1255	10 US-09-811-123-9	Sequence 9, Appl1
15	292	42.3	1255	10 US-09-811-115-3	Sequence 3, Appl1
16	289	41.9	289	10 US-09-821-883-23	Sequence 23, Appl1
17	217	31.4	217	10 US-09-821-883-25	Sequence 25, Appl1
18	217	31.4	397	10 US-09-821-883-27	Sequence 27, Appl1
19	217	31.4	1179	10 US-09-821-883-29	Sequence 29, Appl1

20	173	25.1	191	9 US-09-441-411-9	Sequence 9, Appl1
21	132	19.1	266	9 US-09-854-356-4	Sequence 4, Appl1
22	132	19.1	583	9 US-09-930-125-9	Sequence 9, Appl1
23	132	19.1	587	9 US-09-930-125-8	Sequence 8, Appl1
24	132	19.1	589	9 US-09-930-125-10	Sequence 10, Appl1
25	132	19.1	600	9 US-09-930-125-11	Sequence 11, Appl1
26	127	18.4	127	9 US-09-792-793A-15	Sequence 15, Appl1
27	127	18.4	127	10 US-09-821-883-18	Sequence 18, Appl1
28	127	18.4	127	10 US-09-800-016-1	Sequence 1, Appl1
29	127	18.4	144	10 US-09-923-246-114	Sequence 114, Appl1
30	127	18.4	610	10 US-09-783-708-1	Sequence 1, Appl1
31	117	17.0	135	10 US-09-925-301-1232	Sequence 1232, Appl1
32	52	7.5	654	9 US-09-854-356-8	Sequence 8, Appl1
33	52	7.5	1256	9 US-09-854-356-2	Sequence 2, Appl1
34	52	7.5	1260	9 US-09-870-759-118	Sequence 118, Appl1
35	46	6.7	293	9 US-10-102-806-583	Sequence 583, Appl1
36	44	6.4	1256	9 US-09-854-356-14	Sequence 14, Appl1
37	35	5.1	171	10 US-09-925-300-1025	Sequence 1025, Appl1
38	35	5.1	386	9 US-09-895-793-945	Sequence 945, Appl1
39	35	5.1	386	9 US-09-974-546-48	Sequence 48, Appl1
40	35	5.1	386	10 US-09-822-827-945	Sequence 945, Appl1
41	32	4.6	32	10 US-09-821-883-11	Sequence 11, Appl1
42	22	3.2	22	10 US-09-466-320-19	Sequence 19, Appl1
43	22	3.2	68	10 US-09-466-320-11	Sequence 11, Appl1
44	18	2.6	19	10 US-09-466-320-20	Sequence 20, Appl1
45	18	2.6	65	10 US-09-466-320-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1

US-09-821-883-2

Sequence 2, Application US/09821883

Patent No. US20020061310A1

GENERAL INFORMATION:

APPLICANT: Laus, Retner

APPLICANT: Vidovics, Thomas

APPLICANT: Graddis, Damar

TITLE OF INVENTION: Compositions and Methods for Dendritic

FILE REFERENCE: 7636-0022.30

CURRENT APPLICATION NUMBER: US/09/821, 883

CURRENT FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: US 60/193, 504

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 690

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: HERS500-HGM-CSF construct

US-09-821-883-2

Query Match 100.0%; Score 690; DB 10; Length 690;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRAAPLLARAASLSLGLFLLFWLDRSVLAKELARGAASQVCTGDMKRLRASPET	60
DB	1	MRAAPLLARAASLSLGLFLLFWLDRSVLAKELARGAASQVCTGDMKRLRASPET	60
QY	61	HIDMRLHLYOGGVQGNLELTYLPTNASTSLFDIOEYQGVYLAHNOVRYPLQRLI	120
DB	61	HIDMRLHLYOGGVQGNLELTYLPTNASTSLFDIOEYQGVYLAHNOVRYPLQRLI	120
QY	121	VAGTGFEDNVALVLDNGDPLNNTPTVYASPGGRLQSLTEILKGVLIORNPOL	180
DB	121	VAGTGFEDNVALVLDNGDPLNNTPTVYASPGGRLQSLTEILKGVLIORNPOL	180
QY	181	CYODTILMKDIFHKNNOLATLIDITNRSRACHPCSPMGKGRCSWESSEDCOSLTRTVCA	240

```
Db 181 CYPOTILMKDIFHNKNOTALTLIDTNRSRACHPCSPMKGRCWGESSEDCOSLTRVCA 240
Qy 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALVTYNTDPRES 300
Db 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALVTYNTDPRES 300
Qy 301 MPNEBGRYTFGASCVTACPYNYLSTDVSGAGGAGVNHHRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNEBGRYTFGASCVTACPYNYLSTDVSGAGGAGVNHHRSSSTRSGGDLTLGLEPSEE 360
Qy 361 EAPRSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQORYSEDPVLPSPSTDSYV 420
Db 361 EAPRSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQORYSEDPVLPSPSTDSYV 420
Qy 421 APLTCSPOPEYVNPQDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVADVFAFG 480
Db 421 APLTCSPOPEYVNPQDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVADVFAFG 480
Qy 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPTAENPEY 540
Qy 541 LGLVPAAPARSPSPSTQPPWEHNAIOEARLLNSDPTAEMNTEYEVISEMFDQEP 600
Db 541 LGLVPAAPARSPSPSTQPPWEHNAIOEARLLNSDPTAEMNTEYEVISEMFDQEP 600
Qy 601 TCLQTRLELYKQGLRSLTKLKGPLTMMASHYKOHCPPTSCATQITTESFRENKLD 660
Db 601 TCLQTRLELYKQGLRSLTKLKGPLTMMASHYKOHCPPTSCATQITTESFRENKLD 660
Qy 661 FLVVIPEDCWEPVOGAPPPAAAHNNHH 690
Db 661 FLVVIPEDCWEPVOGAPPPAAAHNNHH 690
```

RESULT 2

US-09-821-883-1

```
: Sequence 1, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Retner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: SOFTWARE: FastSeq for Windows Version 4.0
: NUMBER OF SEQ ID NOS: 30
: SEQ ID NO. 1
: LENGTH: 555
: TYPE: PRN
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500 construct
US-09-821-883-1
```

```
Query Match 79.6%; Score 549; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 MRAAPLLARASLSLGFLLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60
Db 1 MRAAPLLARASLSLGFLLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLHYGCGVYVGNLELTYLPTNASLSFLDODIOEVGYVLIANOVQVPLQRLRI 120
Db 61 HLDMLRHLHYGCGVYVGNLELTYLPTNASLSFLDODIOEVGYVLIANOVQVPLQRLRI 120

```
Qy 121 VRGTQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLREQLRLSTELIKGVLIOBNPOL 180
Db 121 VRGTQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLREQLRLSTELIKGVLIOBNPOL 180
Qy 181 CYPOTILMKDIFHNKNOTALTLIDTNRSRACHPCSPMKGRCWGESSEDCOSLTRVCA 240
Db 181 CYPOTILMKDIFHNKNOTALTLIDTNRSRACHPCSPMKGRCWGESSEDCOSLTRVCA 240
Qy 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALVTYNTDPRES 300
Db 241 GGCARCKGRLPTDCCHEOCAGCTGPKHSDCLACLFHNSGICELHCALVTYNTDPRES 300
Qy 301 MPNEBGRYTFGASCVTACPYNYLSTDVSGAGGAGVNHHRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNEBGRYTFGASCVTACPYNYLSTDVSGAGGAGVNHHRSSSTRSGGDLTLGLEPSEE 360
Qy 361 EAPRSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQORYSEDPVLPSPSTDSYV 420
Db 361 EAPRSPLAPSGAGSDVFDGDLGMAAGLQSLPHDPSLQORYSEDPVLPSPSTDSYV 420
Qy 421 APLTCSPOPEYVNPQDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVADVFAFG 480
Db 421 APLTCSPOPEYVNPQDVAPQPPSPREGPLPAARAGATLEBAKTLSPKNGVADVFAFG 480
Qy 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPAFSPAFDNLYYWDODPPERGAPSTFKGTPTAENPEY 540
Qy 541 LGLVPAAPARSPSPSTQPPWEHNAIOEARLLNSDPTAEMNTEYEVISEMFDQEP 600
Db 541 LGLVPAAPARSPSPSTQPPWEHNAIOEARLLNSDPTAEMNTEYEVISEMFDQEP 600
```

RESULT 3

US-09-821-883-5

```
: Sequence 5, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Retner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: SOFTWARE: FastSeq for Windows Version 4.0
: NUMBER OF SEQ ID NOS: 30
: SEQ ID NO. 5
: LENGTH: 479
: TYPE: PRN
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER300*-TGM-CSF construct
US-09-821-883-5
```

```
Query Match 47.7%; Score 329; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 4,1e-267;
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 MRAAPLLARASLSLGFLLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60
Db 1 MRAAPLLARASLSLGFLLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLHYGCGVYVGNLELTYLPTNASLSFLDODIOEVGYVLIANOVQVPLQRLRI 120
Db 61 HLDMLRHLHYGCGVYVGNLELTYLPTNASLSFLDODIOEVGYVLIANOVQVPLQRLRI 120

Qy 121 VRGTQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLREQLRLSTELIKGVLIOBNPOL 180
Db 121 VRGTQLEFEDNTALAVLNDGDLNNTTPVTGASPGGLREQLRLSTELIKGVLIOBNPOL 180

QY 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMCKSGKSCWESSESDCSLTRTYCA 240
Db 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMCKSGKSCWESSESDCSLTRTYCA 240
QY 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
QY 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329
Db 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329

RESULT 4

US-09-821-883-3
; Sequence 3, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 47.7%; Score 329; DB 10; Length 564;

Best Local Similarity 100.0%; Pred. No. 4.8e-267; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELAGAASSTOVCSTGDMKRLPASPT 60
Db 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELAGAASSTOVCSTGDMKRLPASPT 60
QY 61 HLDMLRHLHYOGCOVVOGNLELTYLPTNASTSLFLODIOEVQGVLLAHNOVROVPLQRLRI 120
Db 61 HLDMLRHLHYOGCOVVOGNLELTYLPTNASTSLFLODIOEVQGVLLAHNOVROVPLQRLRI 120
QY 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGSLRELOLSLEILLKGVLIQRNPOL 180
Db 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGSLRELOLSLEILLKGVLIQRNPOL 180
QY 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMCKSGKSCWESSESDCSLTRTYCA 240
Db 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMCKSGKSCWESSESDCSLTRTYCA 240
QY 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
QY 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329
Db 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329

RESULT 5

US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:

; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-RGM-CSF construct
US-09-821-883-4

Query Match 47.7%; Score 329; DB 10; Length 697;

Best Local Similarity 100.0%; Pred. No. 5.7e-267; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELAGAASSTOVCSTGDMKRLPASPT 60
Db 1 MRAAPLLARASLSLGFLLFFWLDKRSVLAKELAGAASSTOVCSTGDMKRLPASPT 60
QY 61 HLDMLRHLHYOGCOVVOGNLELTYLPTNASTSLFLODIOEVQGVLLAHNOVROVPLQRLRI 120
Db 61 HLDMLRHLHYOGCOVVOGNLELTYLPTNASTSLFLODIOEVQGVLLAHNOVROVPLQRLRI 120
QY 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGSLRELOLSLEILLKGVLIQRNPOL 180
Db 121 VRGTOLEFEDNTALAVLDNGDPLNNTPTVYGASPGSLRELOLSLEILLKGVLIQRNPOL 180
QY 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMCKSGKSCWESSESDCSLTRTYCA 240
Db 181 CYODTILMKDIFHKNNQALATLIDITNRSRACHPCSPMCKSGKSCWESSESDCSLTRTYCA 240
QY 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGRLPTDCCHCCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
QY 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329
Db 301 MPNPGRTFGASCVTACPYNTLSTDVGS 329

RESULT 6

US-09-921-161-1
; Sequence 1, Application US/09921161
; Patent No. US20020090662A1
; GENERAL INFORMATION:
; APPLICANT: Ralph, Peter
; TITLE OF INVENTION: ANALYTICAL METHOD
; FILE REFERENCE: GENENT.066A
; CURRENT APPLICATION NUMBER: US/09/921,161
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/225,433
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-161-1

Query Match 42.3%; Score 292; DB 10; Length 645;

Best Local Similarity 100.0%; Pred. No. 4.7e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 38 GAATGCTGDMKRLPASPETHLDMRLHYOCQVQGNLELYLPTNASLSFLDIO 97
|
|
|
Db 19 GAATGCTGDMKRLPASPETHLDMRLHYOCQVQGNLELYLPTNASLSFLDIO 78
|
|
|
QY 98 EVGGVYLIANQVQVPLQRLIRYRGTOLEFEDNALAVLDGDDPLNNTPTGASPGGLR 157
|
|
|
Db 79 EVGGVYLIANQVQVPLQRLIRYRGTOLEFEDNALAVLDGDDPLNNTPTGASPGGLR 138
|
|
|
QY 158 ELQRLSLTEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSPGGLR 217
|
|
|
Db 139 ELQRLSLTEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSPGGLR 198
|
|
|
QY 218 CKGRRCWGESSEDCOSLTRVCAGGACGKGPLPTDCHEGCAAGCTGPKHSDCLAHF 277
|
|
|
Db 199 CKGRRCWGESSEDCOSLTRVCAGGACGKGPLPTDCHEGCAAGCTGPKHSDCLAHF 258
|
|
|
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
|
|
|
Db 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310
|
|
|
```

```
RESULT 7
: US-09-854-356-3
: Sequence 3, Application US/09854356
: Patent No. US2002017567A1
```

```
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 3
: LENGTH: 653
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
: US-09-854-356-3
```

```
Query Match 42.3%: Score 292; DB 9; Length 653;
```

```
Best Local Similarity 100.0%; Pred. No. 4,8e-236;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 38 GAATGCTGDMKRLPASPETHLDMRLHYOCQVQGNLELYLPTNASLSFLDIO 97
|
|
|
Db 19 GAATGCTGDMKRLPASPETHLDMRLHYOCQVQGNLELYLPTNASLSFLDIO 78
|
|
|
QY 98 EVGGVYLIANQVQVPLQRLIRYRGTOLEFEDNALAVLDGDDPLNNTPTGASPGGLR 157
|
|
|
Db 79 EVGGVYLIANQVQVPLQRLIRYRGTOLEFEDNALAVLDGDDPLNNTPTGASPGGLR 138
|
|
|
QY 158 ELQRLSLTEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSPGGLR 217
|
|
|
Db 139 ELQRLSLTEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSPGGLR 198
|
|
|
QY 218 CKGRRCWGESSEDCOSLTRVCAGGACGKGPLPTDCHEGCAAGCTGPKHSDCLAHF 277
|
|
|
Db 199 CKGRRCWGESSEDCOSLTRVCAGGACGKGPLPTDCHEGCAAGCTGPKHSDCLAHF 258
|
|
|
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
|
|
|
Db 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310
|
|
|
```

```
RESULT 8
: US-09-854-356-7
: Sequence 7, Application US/09854356
: Patent No. US2002017567A1
```

```
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 7
: LENGTH: 712
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: fusion protein
: US-09-854-356-7
```

```
Query Match 42.3%: Score 292; DB 9; Length 712;
```

```
Best Local Similarity 100.0%; Pred. No. 5.1e-236;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 38 GAATGCTGDMKRLPASPETHLDMRLHYOCQVQGNLELYLPTNASLSFLDIO 97
|
|
|
Db 19 GAATGCTGDMKRLPASPETHLDMRLHYOCQVQGNLELYLPTNASLSFLDIO 78
|
|
|
QY 98 EVGGVYLIANQVQVPLQRLIRYRGTOLEFEDNALAVLDGDDPLNNTPTGASPGGLR 157
|
|
|
Db 79 EVGGVYLIANQVQVPLQRLIRYRGTOLEFEDNALAVLDGDDPLNNTPTGASPGGLR 138
|
|
|
QY 158 ELQRLSLTEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSPGGLR 217
|
|
|
Db 139 ELQRLSLTEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRSPGGLR 198
|
|
|
QY 218 CKGRRCWGESSEDCOSLTRVCAGGACGKGPLPTDCHEGCAAGCTGPKHSDCLAHF 277
|
|
|
Db 199 CKGRRCWGESSEDCOSLTRVCAGGACGKGPLPTDCHEGCAAGCTGPKHSDCLAHF 258
|
|
|
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
|
|
|
Db 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310
|
|
|
```

```
RESULT 9
```

```
: US-09-854-356-6
: Sequence 6, Application US/09854356
: Patent No. US2002017567A1
```

```
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 6
```

LENGTH: 919
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE: Description of Artificial Sequence: fusion protein
OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

Query Match 42.3%; Score 292; DB 9; Length 919;
Best Local Similarity 100.0%; Pred. No. 6.4e-236;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASTGVCTGTDKRLPASPETHDMLRHLVGGCOVGNLELYLPTNASTSFLQDIO 97
DB 19 GAASTGVCTGTDKRLPASPETHDMLRHLVGGCOVGNLELYLPTNASTSFLQDIO 78
QY 98 EVQGYVLIHANOYRVPDLRLRVGTQLEFEDNYALAVLDNGDPLNNTTPVYGASPGGLR 157
DB 79 EVQGYVLIHANOYRVPDLRLRVGTQLEFEDNYALAVLDNGDPLNNTTPVYGASPGGLR 138
QY 158 ELQLRSLTEILKGGVLIQNRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQLRSLTEILKGGVLIQNRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
QY 218 CKGSRWGSSESDCOSLRTVCAAGCARGKGLPTDCCHEQCAAGCTGPKHSDDLCLH 277
DB 199 CKGSRWGSSESDCOSLRTVCAAGCARGKGLPTDCCHEQCAAGCTGPKHSDDLCLH 258
QY 278 NHSGICELHCPALVYNTDTFESMPNBEGRYTFGASCVTACPYNYLSTDVGS 329
DB 259 NHSGICELHCPALVYNTDTFESMPNBEGRYTFGASCVTACPYNYLSTDVGS 310

RESULT 10

US-09-769-508-2

Sequence 2, Application US/09769508
Patent No. US20020155527A1
GENERAL INFORMATION:
APPLICANT: STUART, SUSAN G.
APPLICANT: MONAHAN, JOHN J.
APPLICANT: LANGTON, BEATRICE CLAUDIA
APPLICANT: HANCOCK, MIRIAM E.C.
APPLICANT: BLUFORD, PETER
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
FILE REFERENCE: BEBO-111-C1
CURRENT APPLICATION NUMBER: US/09/769,508
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-769-508-2

Query Match 42.3%; Score 292; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 8.4e-236;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASTGVCTGTDKRLPASPETHDMLRHLVGGCOVGNLELYLPTNASTSFLQDIO 97
DB 19 GAASTGVCTGTDKRLPASPETHDMLRHLVGGCOVGNLELYLPTNASTSFLQDIO 78
QY 98 EVQGYVLIHANOYRVPDLRLRVGTQLEFEDNYALAVLDNGDPLNNTTPVYGASPGGLR 157
DB 79 EVQGYVLIHANOYRVPDLRLRVGTQLEFEDNYALAVLDNGDPLNNTTPVYGASPGGLR 138
QY 158 ELQLRSLTEILKGGVLIQNRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQLRSLTEILKGGVLIQNRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
QY 218 CKGSRWGSSESDCOSLRTVCAAGCARGKGLPTDCCHEQCAAGCTGPKHSDDLCLH 277

DB 199 CKGSRWGSSESDCOSLRTVCAAGCARGKGLPTDCCHEQCAAGCTGPKHSDDLCLH 258
QY 278 NHSGICELHCPALVYNTDTFESMPNBEGRYTFGASCVTACPYNYLSTDVGS 329
DB 259 NHSGICELHCPALVYNTDTFESMPNBEGRYTFGASCVTACPYNYLSTDVGS 310

RESULT 11

US-09-854-356-1

Sequence 1, Application US/09854356
Patent No. US20020177567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheyssen, Dirk
APPLICANT: Corixa Corporation
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356
CURRENT FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: US 09/493,480
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: US 60/117,976
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1)..(653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676)..(1255)
OTHER INFORMATION: Intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred.
OTHER INFORMATION: portion (delta PD)
US-09-854-356-1

Query Match 42.3%; Score 292; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 8.4e-236;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASTGVCTGTDKRLPASPETHDMLRHLVGGCOVGNLELYLPTNASTSFLQDIO 97
DB 19 GAASTGVCTGTDKRLPASPETHDMLRHLVGGCOVGNLELYLPTNASTSFLQDIO 78
QY 98 EVQGYVLIHANOYRVPDLRLRVGTQLEFEDNYALAVLDNGDPLNNTTPVYGASPGGLR 157
DB 79 EVQGYVLIHANOYRVPDLRLRVGTQLEFEDNYALAVLDNGDPLNNTTPVYGASPGGLR 138
QY 158 ELQLRSLTEILKGGVLIQNRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQLRSLTEILKGGVLIQNRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
QY 218 CKGSRWGSSESDCOSLRTVCAAGCARGKGLPTDCCHEQCAAGCTGPKHSDDLCLH 277
DB 199 CKGSRWGSSESDCOSLRTVCAAGCARGKGLPTDCCHEQCAAGCTGPKHSDDLCLH 258
QY 278 NHSGICELHCPALVYNTDTFESMPNBEGRYTFGASCVTACPYNYLSTDVGS 329
DB 259 NHSGICELHCPALVYNTDTFESMPNBEGRYTFGASCVTACPYNYLSTDVGS 310

```

RESULT 12
US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Vegdick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-930-125-2

```

```

Query Match          42.3%; Score 292; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 8.4e-236;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 38 GAATGTCGTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTFLODIO 97
DB 19 GAATGTCGTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTFLODIO 78
OY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 157
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138
OY 158 ELQRLSLTEILKGVLIQRLRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGVLIQRLRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198
OY 218 CKGRKMGESSEDCOSLRTVCAGGACARCKGPLPTDCCHBCCAGCTGPKHSDCLACLHF 277
DB 199 CKGRKMGESSEDCOSLRTVCAGGACARCKGPLPTDCCHBCCAGCTGPKHSDCLACLHF 258
OY 278 NMSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

```

```

RESULT 13
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Hellstrom, Ingegerd
; APPLICANT: Hellstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6
Query Match          42.3%; Score 292; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 8.4e-236;

```

```

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 38 GAATGTCGTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTFLODIO 97
DB 19 GAATGTCGTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTFLODIO 78
OY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 157
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138
OY 158 ELQRLSLTEILKGVLIQRLRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGVLIQRLRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198
OY 218 CKGRKMGESSEDCOSLRTVCAGGACARCKGPLPTDCCHBCCAGCTGPKHSDCLACLHF 277
DB 199 CKGRKMGESSEDCOSLRTVCAGGACARCKGPLPTDCCHBCCAGCTGPKHSDCLACLHF 258
OY 278 NMSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

```

```

RESULT 14
US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US2002001587A1
; GENERAL INFORMATION:
; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-ETBB
; TITLE OF INVENTION: ANTIBODY-MAINTAINING CONJUGATES
; FILE REFERENCE: GENE 073A2
; CURRENT APPLICATION NUMBER: US/09/811.123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238,327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602,530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9

```

```

Query Match          42.3%; Score 292; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 8.4e-236;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 38 GAATGTCGTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTFLODIO 97
DB 19 GAATGTCGTGDMKRLPASPETHIDMLRHLYGCGVQVGNLELYLPTNLSLSTFLODIO 78
OY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 157
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138
OY 158 ELQRLSLTEILKGVLIQRLRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGVLIQRLRNPOLCYQDTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198
OY 218 CKGRKMGESSEDCOSLRTVCAGGACARCKGPLPTDCCHBCCAGCTGPKHSDCLACLHF 277
DB 199 CKGRKMGESSEDCOSLRTVCAGGACARCKGPLPTDCCHBCCAGCTGPKHSDCLACLHF 258
OY 278 NMSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

```


RESULT 15
US-09-811-115-3
: Sequence 3, Application US/09811115
: Patent No. US20020035736A1
: GENERAL INFORMATION:
: APPLICANT: Erickson, Sharon
: APPLICANT: Schwall, Ralph
: APPLICANT: King, Kathleen
: TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
: FILE REFERENCE: GENENT 034A
: CURRENT APPLICATION NUMBER: US/09/811,115
: CURRENT FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/189,844
: PRIOR FILING DATE: 2000-03-16
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 1255
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 42.38; Score 292; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 8.4e-236;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 GAATGCTGTDMKRLPASPETHDMLRHLYGCGVVGNNLELTPTNASTSLFDIO 97
DB 19 GAATGCTGTDMKRLPASPETHDMLRHLYGCGVVGNNLELTPTNASTSLFDIO 78
QY 98 EVGGVLLAHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 157
DB 79 EVGGVLLAHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVTGASPGGLR 138
QY 158 ELQLRSLLEILKGGVLIQRNQLCYQDTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQLRSLLEILKGGVLIQRNQLCYQDTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
QY 218 CKGRMCWSESSEDCSLTRTYCAGGACARCKGPLPTDCCHQCAAGCTGPKHSDCLACLHF 277
DB 199 CKGRMCWSESSEDCSLTRTYCAGGACARCKGPLPTDCCHQCAAGCTGPKHSDCLACLHF 258
QY 278 NMSGICELHCPALVTVNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVTVNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 310

Search completed: April 28, 2003, 13:47:06
Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:08 ; Search time 23.1973 Seconds

(without alignments)
1660.081 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGVCTGTDKRLRLPASPT.....FGASCVACPYNTLSTDVGS 289

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	289	AAE13120	Mature human HER-2
2	1587	100.0	479	AAE13112	Human HER300-TCM-C
3	1587	100.0	535	AAE13108	Human HER500 fusio
4	1587	100.0	564	AAE13110	Human HER500 oncosp
5	1587	100.0	645	AAE13108	Human HER500 oncosp
6	1587	100.0	645	AAE13108	Human HER500 oncosp
7	1587	100.0	645	AAE13108	Human HER500 oncosp
8	1587	100.0	645	AAE13108	Human HER500 oncosp
9	1587	100.0	645	AAE13108	Human HER500 oncosp
10	1587	100.0	645	AAE13108	Human HER500 oncosp

11	1587	100.0	712	21	AAE13120	Human HER-2/neu fu
12	1587	100.0	712	23	AAE13149	Her-2/neu extracel
13	1587	100.0	782	18	AAE19764	Her-2/neu extracel
14	1587	100.0	919	21	AAE13120	Human HER-2/neu fu
15	1587	100.0	919	21	AAE13120	Human HER-2/neu fu
16	1587	100.0	951	21	AAE14993	Her-2/neu extracel
17	1587	100.0	1200	21	AAE13120	Human HER-2/neu fu
18	1587	100.0	1255	17	AAE13120	Human HER-2/neu fu
19	1587	100.0	1255	20	AAE13120	Human HER-2/neu fu
20	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
21	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
22	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
23	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
24	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
25	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
26	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
27	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
28	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
29	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
30	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
31	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
32	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
33	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
34	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
35	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
36	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
37	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
38	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
39	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
40	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
41	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
42	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
43	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
44	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu
45	1587	100.0	1255	21	AAE13120	Human HER-2/neu fu

ALIGNMENTS

RESULT 1	AAE13120	standard; Protein: 289 AA.
ID	AAE13120	
AC	AAE13120	
DT	28-JAN-2002	(first entry)
XX		
DE	Mature human HER-2 membrane distal extracellular domain.	
XX		
KW	Immunostimulatory fusion protein; IFP; antigen component; therapy;	
KW	immunostimulatory component; T-cell mediated immune response; DC;	
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;	
KW	human; HER-2 membrane distal extracellular domain.	
OS	Homo sapiens.	
XX		
PN	WO200174855 A2.	
XX		
PD	11-OCT-2001.	
XX		
PF	30-MAR-2001; 2001WO-US10515.	
XX		
PR	30-MAR-2000; 2000US-193504P.	
XX		
PA	(DEND-) DENDREON CORP.	
XX		
PI	Laus R, Vidovic D, Graddis T;	
XX		
DR	WPI: 2001-662965/76.	
XX		
PT	N-PSDB: AAD21571.	
XX		
XX	An immunostimulatory fusion protein comprising the intracellular domain	

PT of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 PS Claim 6; Page 30; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is
 CC mature human HER-2 membrane distal extracellular domain.
 CC This sequence is used in the HER500 and HER300 GM-CSF fusion constructs
 CC of the invention.

XX Sequence 289 AA:

Query Match 100.0%; Score 1587; DB 22; Length 289;
 Best Local Similarity 100.0%; Pred. No. 7,8e-132;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASPERHDLRLHLYGCGVVGNNLELTYPYNASLSFLDDIDDEVQ 60
 DB 1 STQVCTGDMKRLRPASPERHDLRLHLYGCGVVGNNLELTYPYNASLSFLDDIDDEVQ 60
 QY 61 GYVLIANNOYRQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELO 120
 DB 61 GYVLIANNOYRQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELO 120
 QY 121 LRSLTLEIKGVLIOIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 180
 DB 121 LRSLTLEIKGVLIOIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 180
 QY 181 SRCWGESSEDCQSLTRIVCAGGCARCKGPLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 240
 DB 181 SRCWGESSEDCQSLTRIVCAGGCARCKGPLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 240
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 289
 DB 241 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 289

RESULT 2

AEI3112 standard; Protein; 479 AA.

XX AEI3112:

XX 28-JAN-2002 (first entry)

DE Human HER300-rGM-CSF fusion construct comprising OVA-derived peptide.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW Immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
 KW ovalbumin-derived octapeptide; OVA; rat; HER300-rGM-CSF fusion protein.

XX Chimeric - Homo sapiens.

OS Chimeric - Rattus norvegicus.

XX Chimeric - Unidentified.

XX WO200174855-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001MO-US10515.
 XX 30-MAR-2000; 2000US-193504P.

XX (DEND-) DENDREON CORP.
 PA Laus: R, Vidovic D, Graddis T;
 XX WPI; 2001-662965/76.
 DR N-PDB; AAD21568.

PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 PS Example 1; Page 27; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER300
 CC rGM-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC signal sequence, mature HER-2 membrane distal extracellular domain,
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
 CC a rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
 CC and a C-terminal tag.

XX Sequence 479 AA:

Query Match 100.0%; Score 1587; DB 22; Length 479;
 Best Local Similarity 100.0%; Pred. No. 1.5e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKRLRPASPERHDLRLHLYGCGVVGNNLELTYPYNASLSFLDDIDDEVQ 60
 DB 41 STQVCTGDMKRLRPASPERHDLRLHLYGCGVVGNNLELTYPYNASLSFLDDIDDEVQ 100
 QY 61 GYVLIANNOYRQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELO 120
 DB 101 GYVLIANNOYRQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLNNTTPTVTGASPGGLRELO 160
 QY 121 LRSLTLEIKGVLIOIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 180
 DB 161 LRSLTLEIKGVLIOIRNPOLCYODTILMKDIFRKNQALATLIDTNSRACHPCSPCKG 220
 QY 181 SRCWGESSEDCQSLTRIVCAGGCARCKGPLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 240
 DB 221 SRCWGESSEDCQSLTRIVCAGGCARCKGPLPTDCCHQCAAGCTGPRHSDCLACLHFNHS 280
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 289
 DB 281 GICELHCPALVTYNTDFESMPNPEGRTYTGASCVTACPNYISTDVGS 329

RESULT 3

AEI3108 standard; Protein; 555 AA.

XX AEI3108:

XX 28-JAN-2002 (first entry)

DE Human HER500 fusion protein construct.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW Immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human;
 KW HER-2 protein; HER500 fusion protein.

XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Synthetic.
XX PN WO200174855-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US10515.
XX PR 30-MAR-2000; 2000US-193504P.
XX PA (DEND-) DENDREON CORP.
XX PI Laus R, Vidovic D, Graddis T;
XX DR WPI; 2001-662965/76.
XX DR N-PSDB; AAD21564.
XX PT An immunostimulatory fusion protein comprising the intracellular domain
XX PT of HER-2 and an antigen elicits an immune response to the antigen and
XX PT is useful for the treatment of associated cancer associated -
XX PS Claim 7; Page 26; 59pp; English.
XX CC The invention relates to immunostimulatory fusion proteins (IFP) and
XX CC nucleic acid molecules encoding such proteins. The IFPs comprise a
XX CC polypeptide antigen component and an immunostimulatory component derived
XX CC from the intracellular domain of HER-2 protein which is effective to
XX CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX CC immune response to the antigen. IFP or superactivated dendritic cells
XX CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX CC associated with a particularly antigen. The present sequence is HER500
XX CC fusion protein construct which comprises human PAP signal
XX CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
XX CC sequence, mature HER-2 membrane distal extracellular and intracellular
XX CC domains and a C-terminal tag.
XX SQ Sequence 555 AA;
Query Match 100.0%; Score 1587; DB 22; Length 555;
Best Local Similarity 100.0%; Pred. No. 1,8e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 60
DB 41 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 100
OY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 120
DB 101 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 160
OY 121 LRSLEILKGGVLIQORNPOLCYOTIIMKDIFFHKNQALATLIDITNRSRACHPCSPMKG 180
DB 161 LRSLEILKGGVLIQORNPOLCYOTIIMKDIFFHKNQALATLIDITNRSRACHPCSPMKG 220
OY 181 SRCWGESSEDCQSLRTFYCAGCARGKGPLPTDCHECCAGCTGPKHSDCLACLHFHNS 240
DB 221 SRCWGESSEDCQSLRTFYCAGCARGKGPLPTDCHECCAGCTGPKHSDCLACLHFHNS 280
OY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYVYLSIDVGS 289
DB 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYVYLSIDVGS 329
RESULT 4
AAE13110
ID AAE13110 standard; Protein: 564 AA.
XX AAE13110;
XX DT 28-JAN-2002 (first entry)
XX

DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.
XX
XX
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; OVA;
KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein...
XX
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Unidentified.
XX PN WO200174855-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US10515.
XX PR 30-MAR-2000; 2000US-193504P.
XX PA (DEND-) DENDREON CORP.
XX PI Laus R, Vidovic D, Graddis T;
XX DR WPI; 2001-662965/76.
XX DR N-PSDB; AAD21566.
XX PT An immunostimulatory fusion protein comprising the intracellular domain
XX PT of HER-2 and an antigen elicits an immune response to the antigen and
XX PT is useful for the treatment of associated cancer associated -
XX PS Claim 7; Page 26; 59pp; English.
XX CC The invention relates to immunostimulatory fusion proteins (IFP) and
XX CC nucleic acid molecules encoding such proteins. The IFPs comprise a
XX CC polypeptide antigen component and an immunostimulatory component derived
XX CC from the intracellular domain of HER-2 protein which is effective to
XX CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX CC immune response to the antigen. IFP or superactivated dendritic cells
XX CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX CC associated with a particularly antigen. The present sequence is HER500
XX CC fusion protein construct which comprises human PAP signal
XX CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
XX CC sequence, mature HER-2 membrane distal extracellular domain, an
XX CC Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
XX CC HER-2 membrane distal intracellular domain and a C-terminal tag.
XX SQ Sequence 564 AA;
Query Match 100.0%; Score 1587; DB 22; Length 564;
Best Local Similarity 100.0%; Pred. No. 1,9e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 60
DB 41 STQVCTGDMKRLRPASBETHLDMRLHYGCGVVGNGLELTYLPTNASLSFLDIOEVQ 100
OY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 120
DB 101 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 160
OY 121 LRSLEILKGGVLIQORNPOLCYOTIIMKDIFFHKNQALATLIDITNRSRACHPCSPMKG 180
DB 161 LRSLEILKGGVLIQORNPOLCYOTIIMKDIFFHKNQALATLIDITNRSRACHPCSPMKG 220
OY 181 SRCWGESSEDCQSLRTFYCAGCARGKGPLPTDCHECCAGCTGPKHSDCLACLHFHNS 240
DB 221 SRCWGESSEDCQSLRTFYCAGCARGKGPLPTDCHECCAGCTGPKHSDCLACLHFHNS 280
OY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYVYLSIDVGS 289
DB 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYVYLSIDVGS 329

```

RESULT 5
AAB60408
ID AAB60408 standard; Protein; 645 AA.
XX
AC AAB60408;
XX
DT 24-APR-2001 (first entry)
XX
DE Human ErbB2 oncoprotein, SEQ ID NO:13.
XX
KW Anti-ErbB2 monoclonal antibody 2C4; HER2; mouse; murine; humanised; VL:
KW light chain variable region; cancer; cytostatic; EGFR-expressing cancer;
KW epidermal growth factor receptor; colon cancer; rectal cancer; tumour;
KW colorectal cancer; non-small cell lung cancer; metastatic breast cancer;
KW affinity purification.
XX
OS Homo sapiens.
XX
PN WO200100245-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-US17366.
XX
PR 25-JUN-1999; 99US-0141316.
XX
PA (GETH ) GENENTECH INC.
XX
PI Adams CW, Presta LG, Sliwkowsky M;
XX
DR WPI: 2001-080862/09.
XX
PT Treating cancer in a human, where the cancer expresses epidermal growth
PT factor receptor (EGFR), comprises administering an antibody which binds
PT ErbB2.
XX
PS Example 1; Fig 1A; 89pp; English.
XX
CC The invention relates to a method for treating cancer in a human patient,
CC wherein the cancer expresses epidermal growth factor receptor (EGFR),
CC comprising administering an antibody which binds ErbB2 (HER2; AAB60408).
CC In particular, the anti-ErbB2 antibody is the murine monoclonal antibody
CC 2C4 (AAB60396, AAB60397) or a humanised version of 2C4 (AAB60398),
CC encoding a humanised ErbB2-binding antibody; vectors and host cells
CC comprising such nucleic acids; the recombinant production of a humanised
CC ErbB2-binding antibody; and an immunconjugate comprising a humanised
CC ErbB2-binding antibody and a cytotoxic drug. The ErbB2-binding antibodies
CC act by antagonising ErbB receptors, and as inhibitors of transforming
CC growth factor alpha (TGF-alpha)-activated mitogen activated protein
CC kinase (MAPK). The method of the invention is used for treating cancer,
CC especially colon cancer, rectal cancer, colorectal cancer, lung cancer,
CC (especially non-small cell lung cancer), or breast cancer (especially
CC metastatic breast cancer). The antibodies may also have non-therapeutic
CC uses e.g., as affinity purification agents. Using an antibody which binds
CC to ErbB2 to treat cancer is preferable to the use of EGFR-targeted
CC drugs, as EGFR is also highly expressed in other tissues such as the
CC liver and skin, where the active drug will also bind, with skin toxicity.
CC having been observed for EGFR-targeted drugs. Antibodies which bind
CC ErbB2 are anticipated to have a better safety profile than such drugs.
CC The present sequence represents human ErbB2.
XX
SQ Sequence 645 AA;

```

```

Query Match 100.0%; Score 1587; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STGVCTGTDKRLRPASPEHDLMLRHLVGGCVVGNLELTLYLPNASTLSPFDIOEYQ 60
DB 22 STGVCTGTDKRLRPASPEHDLMLRHLVGGCVVGNLELTLYLPNASTLSPFDIOEYQ 81

```

```

QY 61 GYVLIHNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPYVGASRGILREIQ 120
DB 82 GYVLIHNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPYVGASRGILREIQ 141
QY 121 LRSLETLKGGVLIQRNPOLCYODTILMKDIFPKNNOLATLLIDTNRSRACHCSPMCKG 180
DB 142 LRSLETLKGGVLIQRNPOLCYODTILMKDIFPKNNOLATLLIDTNRSRACHCSPMCKG 201
QY 181 SRCMGESSEDCQSLTRTVACGACGACRCKGRLPTDCHEQCAAGCTGPKHSDCLAFHNHS 240
DB 202 SRCMGESSEDCQSLTRTVACGACGACRCKGRLPTDCHEQCAAGCTGPKHSDCLAFHNHS 261
QY 241 GIEELHCPALVYNTDTFESMPNDEGRYRTGASCYVACPNYSTDVGS 289
DB 262 GIEELHCPALVYNTDTFESMPNDEGRYRTGASCYVACPNYSTDVGS 310

```

```

RESULT 6
AAB61593
ID AAB61593 standard; Protein; 645 AA.
XX
AC AAB61593;
XX
DT 04-APR-2001 (first entry)
XX
DE Human ErbB2 extracellular domain.
XX
KW Human; ErbB2; cytostatic; prostate cancer; receptor tyrosine kinase;
KW antibody; ErbB receptor; monoclonal antibody 2C4; variable light chain.
XX
OS Homo sapiens.
XX
PN WO200100238-A1.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-US17423.
XX
PR 25-JUN-1999; 99US-0141315.
XX
PA (GETH ) GENENTECH INC.
PA (SLOK ) SLOAN KETTERING INST CANCER RES.
PI Agus DB, Scher HT, Sliwkowski MX;
XX
DR WPI: 2001-159131/16.
XX
PT Treating prostate cancer in a human comprises administering an antibody
PT which binds ErbB2 and blocks ligand activation of an ErbB receptor.
XX
PS Disclosure; Fig 1; 93pp; English.
XX
CC The ErbB family of receptor tyrosine kinases are important mediators of
CC cell growth, differentiation and survival. The receptor family includes
CC four distinct members including Epidermal Growth Factor Receptor (EGFR or
CC ErbB1), HER2 (ErbB2 or p185neu), HER3 (ErbB3) and Her4 (ErbB4 or tyro2).
CC The present invention relates to a method for treating prostate cancer.
CC The method comprises administering an antibody which binds ErbB2 and
CC blocks ligand activation of an ErbB receptor. Preferably, the antibody
CC blocks binding of monoclonal antibody 2C4 to ErbB2 and/or blocks
CC TGF-alpha activation of mitogen-activated protein kinase (MAPK). The
CC present sequence is the extracellular domain of human ErbB2.
XX
SQ Sequence 645 AA;

```

```

Query Match 100.0%; Score 1587; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STGVCTGTDKRLRPASPEHDLMLRHLVGGCVVGNLELTLYLPNASTLSPFDIOEYQ 60
DB 22 STGVCTGTDKRLRPASPEHDLMLRHLVGGCVVGNLELTLYLPNASTLSPFDIOEYQ 81

```

QY 61 GYVLIANOVQVPLQRLIRVGTOLFEDNVALVNDGDELNTPTVTGASPGGLRELQ 120
 DB 82 GYVLIANOVQVPLQRLIRVGTOLFEDNVALVNDGDELNTPTVTGASPGGLRELQ 141
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNSRACHPCSPMKG 180
 DB 142 LRSLEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNSRACHPCSPMKG 201
 QY 181 SRCWSESEDDQSLRTVYACAGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
 DB 202 SRCWSESEDDQSLRTVYACAGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
 QY 241 GICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLSTDVGS 289
 DB 262 GICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 7
 AAB21200
 ID AAB21200 standard; protein: 653 AA.
 AC AAB21200;
 DT 12-JAN-2001 (first entry)
 DE Extracellular HER-2/neu protein.
 KW HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
 KW colon cancer.
 OS Unidentified.
 PN WO200044899-A1.
 PD 03-AUG-2000.
 PF 28-JAN-2000; 2000WO-US02164.
 PR 29-JAN-1999; 99US-0117976.
 PA (CORI-) CORIXA CORP.
 PA (SMIK) SMITHKLINE BEECHAM.
 PI Cheever MA, Gheysen D;
 DR WPI: 2000-505976/45.
 DR HER-2/neu extracellular domain/phosphorylation domain fusion proteins
 PT useful for vaccinating against breast, ovarian, colon, lung and
 PT prostate cancers -
 PS Claim 2: Fig 9; 128pp; English.
 CC The present sequence is the extracellular HER-2/neu protein. HER-2/neu is
 CC a member of the tyrosine kinase family of receptor-like glycoproteins and
 CC shows homology to the epidermal growth factor receptor (EGFR). It
 CC probably plays a part in cell growth and/or differentiation. The
 CC HER-2/neu gene is an oncogene. An HER-2/neu fusion protein comprising a
 CC domain may be used to treat or prevent cancer by eliciting or
 CC enhancing an immune response to the HER-2/neu protein. It may be used
 CC to treat malignancies such as breast, ovarian, colon, lung and
 CC prostate cancers, and may be used as an antigen to vaccinate against
 CC these neoplasias.
 SQ Sequence 653 AA;
 Query Match 100.0%; Score 1587; DB 21; Length 653;
 Best Local Similarity 100.0%; Pred. NO.2.2e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STOVCTGDMKRLRPASPEFHLDMLRHLXOCQOVQGNLELTPTNASLSFLQDIOEVQ 60

DB 22 STOVCTGDMKRLRPASPEFHLDMLRHLXOCQOVQGNLELTPTNASLSFLQDIOEVQ 81
 QY 61 GYVLIANOVQVPLQRLIRVGTOLFEDNVALVNDGDELNTPTVTGASPGGLRELQ 120
 DB 82 GYVLIANOVQVPLQRLIRVGTOLFEDNVALVNDGDELNTPTVTGASPGGLRELQ 141
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNSRACHPCSPMKG 180
 DB 142 LRSLEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNSRACHPCSPMKG 201
 QY 181 SRCWSESEDDQSLRTVYACAGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
 DB 202 SRCWSESEDDQSLRTVYACAGCARCKGRLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
 QY 241 GICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLSTDVGS 289
 DB 262 GICELHCPALVTYNTDFESMPNEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 8
 AAM51145
 ID AAM51145 standard; protein: 653 AA.
 AC AAM51145;
 DT 17-JUN-2002 (first entry)
 DE Human Her-2/neu oncoprotein extracellular domain.
 KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 KW tyrosine kinase; receptor; c-erbB2; gene therapy.
 OS Homo sapiens.
 PN WO200212341-A2.
 PD 14-FEB-2002.
 PF 03-AUG-2001; 2001WO-US24283.
 PR 03-AUG-2000; 2000US-0632507.
 PA (CORI-) CORIXA CORP.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Cheever MA, Gheysen D;
 DR WPI: 2002-241743/29.
 DR Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain -
 PS Claim 2: Fig 9; 141pp; English.
 CC The present sequence is that of the extracellular domain of
 CC human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic
 CC self-protein and target for anti-cancer vaccines. The Her-2/neu
 CC gene is amplified and p185 is overexpressed in a variety of cancers,
 CC including breast, ovarian, colon, lung and prostate cancer.
 CC Her-2/neu (see AAM51143) is a member of the tyrosine kinase family
 CC of receptor-like glycoproteins. It comprises an extracellular
 CC domain with homology to the epidermal growth factor receptor
 CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal
 CC intracellular domain that also shows homology to EGFR. Its
 CC overexpression correlates with a poor prognosis in breast and
 CC ovarian cancers. The invention provides Her-2/neu fusion
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines
 CC comprising the fusion proteins or nucleic acid molecules. In
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu
 CC protein is fused to a Her-2/neu intracellular domain or

CC phosphorylation domain (or its DeltapD fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal *ex vivo* with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.

XX Sequence 653 AA:

Query Match 100.0%; Score 1587; DB 23; Length 653;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELYLPYNASLSFLQDIOEVQ 60
 DB 22 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELYLPYNASLSFLQDIOEVQ 81
 QY 61 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVYGASFGGLREIQ 120
 DB 82 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVYGASFGGLREIQ 141
 QY 121 LRSLETLKGGVLIQRNPOLCYODTILMKDIFRKNOLATLTDITNRSRACHCSPCKG 180
 DB 142 LRSLETLKGGVLIQRNPOLCYODTILMKDIFRKNOLATLTDITNRSRACHCSPCKG 201
 QY 181 SRCWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 240
 DB 202 SRCWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 261
 QY 241 GICELHCPALVYNTDTFESMPNREGRYTGGASCYACAPNYLSTDVGS 289
 DB 262 GICELHCPALVYNTDTFESMPNREGRYTGGASCYACAPNYLSTDVGS 310

RESULT 9
 AAE13109

ID AAE13109 standard; protein; 690 AA.

XX AAE13109;

DT 28-JAN-2002 (first entry)

DE Human HER500-hGM-CSF fusion protein construct.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW Immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
 KW HER500-hGM-CSF fusion protein.

XX Chimeric - Homo sapiens.

OS Chimeric - Synthetic.

XX MO200174855-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US10515.

XX 30-MAR-2000; 2000US-193504P.

XX (DEND-) DENDREON CORP.

XX Laus R, Vldovic D, Graddis T;

XX WPI; 2001-662965/76.

DR N-PSDB; AAD21565.

XX An immunostimulatory fusion protein comprising the intracellular domain
 PT of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 XX Claim 7; Page 26; 59pp; English.

XX The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC hGM-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC signal sequence, mature HER-2 membrane distal extracellular and
 CC intracellular domains, an Ala Ala linker, a mature human granulocyte-
 CC macrophage colony stimulating factor (GM-CSF) sequence and a
 CC C-terminal tag.

XX Sequence 690 AA:

Query Match 100.0%; Score 1587; DB 22; Length 690;
 Best Local Similarity 100.0%; Pred. No. 2, 4e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELYLPYNASLSFLQDIOEVQ 60
 DB 41 STQVCTGTDKRLRLPASPEHDLMLRLHYGCGVVGNGLELYLPYNASLSFLQDIOEVQ 100
 QY 61 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVYGASFGGLREIQ 120
 DB 101 GYVLIAHNOVROYPLRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVYGASFGGLREIQ 160
 QY 121 LRSLETLKGGVLIQRNPOLCYODTILMKDIFRKNOLATLTDITNRSRACHCSPCKG 180
 DB 161 LRSLETLKGGVLIQRNPOLCYODTILMKDIFRKNOLATLTDITNRSRACHCSPCKG 220
 QY 181 SRCWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 240
 DB 221 SRCWGESSEDCOSLTRIVCAGGCARCKGPLPTCCHEQCAAGCTGPKHSCLCLHFNHS 280
 QY 241 GICELHCPALVYNTDTFESMPNREGRYTGGASCYACAPNYLSTDVGS 289
 DB 281 GICELHCPALVYNTDTFESMPNREGRYTGGASCYACAPNYLSTDVGS 329

RESULT 10
 AAE13111

ID AAE13111 standard; protein; 697 AA.

XX AAE13111;

DT 28-JAN-2002 (first entry)

DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW Immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
 KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.

XX Chimeric - Homo sapiens.

OS Chimeric - Rattus norvegicus.

XX Chimeric - Unidentified.

XX MO200174855-A2.

XX 11-OCT-2001.
 PD 30-MAR-2001; 2001MO-US10515.
 XX 30-MAR-2000; 2000US-193504P.
 PR 30-MAR-2000; 2000US-193504P.
 XX (DEND-) DENDREON CORP.
 PA (DEND-) DENDREON CORP.
 XX Laus R, Vidovic D, Graddis T;
 PI WPI; 2001-662965/76.
 DR N-PSDB; AAD21567.
 XX An immunostimulatory fusion protein comprising the intracellular domain
 PT of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 XX Claim 7; Page 27; 59pp; English.
 PS The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC rcm-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC signal sequence, mature HER-2 membrane distal extracellular domain,
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
 CC HER-2 membrane distal intracellular domain, an Ala linker, a mature
 CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
 CC and a C-terminal tag.
 CC
 XX Sequence 697 AA;
 SQ Query Match 100.0%; Score 1587; DB 22; Length 697;
 Best Local Similarity 100.0%; Pred. No. 2.4e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STQVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVGNLELTLYPTNLSLFLDIOEVO 60
 DB 41 STQVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVGNLELTLYPTNLSLFLDIOEVO 100
 QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 120
 DB 101 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 160
 QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180
 DB 161 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 220
 QY 181 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 240
 DB 221 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 280
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 289
 DB 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 329
 RESULT 11
 AAB21204
 ID AAB21204 standard; protein: 712 AA.
 XX
 AC AAB21204;
 XX
 DT 12-JAN-2001 (first entry)
 XX
 DE Human HER-2/neu fusion protein.
 XX

KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
 KW colon cancer; fusion protein.
 XX Homo sapiens.
 OS Synthetic.
 XX WO200044899-A1.
 XX
 PD 03-AUG-2000.
 XX
 XX 28-JAN-2000; 2000MO-US02164.
 PF 29-JAN-1999; 99US-0117976.
 PR 29-JAN-1999; 99US-0117976.
 XX
 XX (CORI-) CORIXA CORP.
 PA (SMIT) SMITHKLINE BEECHAM.
 XX
 XX Cheever MA, Gheysen D;
 PI WPI; 2000-505976/45.
 DR HER-2/neu extracellular domain/phosphorylation domain fusion proteins
 XX useful for vaccinating against breast, ovarian, colon, lung and
 XX prostate cancers -
 PS Claim 27; Fig 13; 128pp; English.
 CC The present sequence is a fusion protein comprising the extracellular
 CC domain and a preferred portion of the phosphorylation domain of the human
 CC HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of
 CC receptor-like glycoproteins and shows homology to the epidermal growth
 CC factor receptor (EGFR). It probably plays a part in cell growth and/or
 CC differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion
 CC proteins may be used to treat or prevent cancer by eliciting or enhancing
 CC an immune response to the HER-2/neu protein. They may be used to treat
 CC malignancies such as breast, ovarian, colon, lung and prostate cancers,
 CC and may be used as an antigen to vaccinate against these neoplasias.
 CC
 XX Sequence 712 AA;
 SQ Query Match 100.0%; Score 1587; DB 21; Length 712;
 Best Local Similarity 100.0%; Pred. No. 2.5e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 STQVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVGNLELTLYPTNLSLFLDIOEVO 60
 DB 22 STQVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVGNLELTLYPTNLSLFLDIOEVO 81
 QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 120
 DB 82 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLREIQ 141
 QY 121 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180
 DB 142 LRSLEILKGGVLIQORNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 201
 QY 181 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 240
 DB 202 SRCWGESSEDCQSLRTVYACAGCARGKPLPTDCHECCAGCTGPKHSDCLACHFNHS 261
 QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 289
 DB 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCYTACPYNYLSTDVGS 310
 RESULT 12
 AAM51149
 ID AAM51149 standard; protein: 712 AA.
 XX
 AC AAM51149;
 XX
 DT 17-JUN-2002 (first entry)
 XX

XX Her-2/neu extracellular domain-delta-phosphorylation domain fusion.
 DE Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 XX tyrosine kinase; receptor; c-erbB2; gene therapy.
 KM
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Domain 1..653
 FT /note= "extracellular domain"
 FT Domain 654..712
 FT /note= "phosphorylation domain fragment"
 XX
 PN WO200213341-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 03-AUG-2001; 2001WO-US24283.
 XX
 PR 03-AUG-2000; 2000US-0632507.
 XX
 PA (CORI-) CORIXA CORP.
 XX (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI Cheever MA, Gheysen D;
 XX
 DR WPI: 2002-241743/29.
 XX
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain -
 XX
 PS Claim 37; Fig 13; 141pp; English.
 XX
 CC The present sequence is that of a fusion protein between the
 CC extracellular domain and a fragment (delta) of the phosphorylation
 CC domain of human Her-2/neu (see AAM5113), an oncogenic self-protein
 CC and target for anti-cancer vaccines. The fusion protein can be
 CC obtained by recombinant DNA methods. Her-2/neu overexpression
 CC correlates with a poor prognosis in breast and ovarian cancers.
 CC The invention provides Her-2/neu fusion proteins, nucleic acids
 CC encoding them, viral vectors, and vaccines comprising the fusion
 CC proteins or nucleic acid molecules. In preferred fusion proteins,
 CC the extracellular domain of Her-2/neu is fused to a Her-2/neu
 CC intracellular domain or phosphorylation domain (or its Delta) or
 CC fragment). An immune response to Her-2/neu protein is elicited or
 CC enhanced by administering the fusion protein in the form of a vaccine,
 CC or by transfecting cells of an animal *ex vivo* with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.
 XX
 SQ Sequence 712 AA:
 XX
 Query Match 100.0%; Score 1587; DB 23; Length 712;
 Best Local Similarity 100.0%; Pred. No. 2.5e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 STQVCTGTDMLKRLPASPETHDMLRHLVGGCCVGGNLETLPLTNASTSLFDQIDQEVQ 60
 DB 22 STQVCTGTDMLKRLPASPETHDMLRHLVGGCCVGGNLETLPLTNASTSLFDQIDQEVQ 81
 OY 61 GYVLIAHNOVROVPLDRLRIVRGTQLFEDNYALAVLDNCDPLNTTPVTGASFGGLREIQ 120
 DB 82 GYVLIAHNOVROVPLDRLRIVRGTQLFEDNYALAVLDNCDPLNTTPVTGASFGGLREIQ 141
 OY 121 LRSLETLKGVLIQRNPOLCYODTILMKDIFKNNQALTLIDTNRSRACHCSPCKG 180

DB 142 LRSLETLKGVLIQRNPOLCYODTILMKDIFKNNQALTLIDTNRSRACHCSPCKG 201
 OY 181 SRCMGESSEDCOSLTRVCGAGCARGKPLPTQCHBOCAAGCTGPKHSPCLACLFHNHS 240
 DB 202 SRCMGESSEDCOSLTRVCGAGCARGKPLPTQCHBOCAAGCTGPKHSPCLACLFHNHS 241
 OY 241 GICELHCPALVTYNTDFESMPNDEGRYTGASCVCYACPNYLTSDVGS 289
 DB 262 GICELHCPALVTYNTDFESMPNDEGRYTGASCVCYACPNYLTSDVGS 310
 RESULT 13
 AAM19764
 ID AAM19764 standard; Protein; 782 AA.
 XX
 AC AAM19764;
 XX
 DT 17-SEP-1997 (first entry)
 XX
 DE Her-2-GM-CSF immunostimulant fusion protein.
 XX
 KM Her-2-GM-CSF; granulocyte macrophage colony stimulating factor;
 KM growth factor receptor; oncogene; immunostimulant; cancer;
 XX therapy.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Protein 1..653
 FT Peptide /label= Her2
 FT /label= 654..655
 FT Protein /label= Linker
 FT /label= 656..782
 FT /label= GM-CSF
 XX
 PN MO9724438-A1.
 XX
 PD 10-JUL-1997.
 XX
 PF 23-DEC-1996; 96WO-US20241.
 XX
 PR 28-DEC-1995; 95US-0579823.
 XX
 PA (ACTI-) ACTIVATED CELL THERAPY INC.
 XX
 PI Laus R, Ruegg CL, Wu H;
 XX
 DR WPI: 1997-363674/33.
 XX N-PSDB; AAT72725.
 XX
 PT Potent APC that activates T-cells to give multivalent cellular
 PT immune response - can also induce a cytotoxic T-cell response in a
 PT vertebrate subject
 XX
 PS Disclosure; Fig 8; 45pp; English.
 XX
 CC A fusion protein (AAM19764) comprises Her2 (a growth factor receptor
 CC that is overexpressed in breast, ovarian and other cancer cells)
 CC and granulocyte-macrophage colony stimulating factor (GM-CSF). It
 CC is the expression product of a nucleic acid molecule (AAT72725)
 CC prep. by PCR amplification of Her2 cDNA from a breast cancer cell
 CC line and fusion to GM-CSF cDNA. Fusion expression vectors can be
 CC used to transfect mammalian and insect cells. The Her2-GM-CSF
 CC fusion protein is used to generate anti-Her2 immunity. Tumour
 CC cells are eliminated by cytotoxic T lymphocytes activated *in vivo*
 CC or *in vitro* by exposure to antigen-presenting cells exposed to the
 CC fusion protein.
 XX
 SQ Sequence 782 AA:
 XX
 Query Match 100.0%; Score 1587; DB 18; Length 782;
 Best Local Similarity 100.0%; Pred. No. 2.8e-131;

Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGYCTGDMKRLRPASPEETHLDMRLHLYGCGVVOGNGLELTYPPTNASLSFLDIOEVQ 60
D 22 STGYCTGDMKRLRPASPEETHLDMRLHLYGCGVVOGNGLELTYPPTNASLSFLDIOEVQ 81
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPTGASPGGLRELQ 120
D 82 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPTGASPGGLRELQ 141
QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRACHPSPCKG 180
D 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRACHPSPCKG 201
QY 181 SRCWGESSEDCQSLTRTYCAGGACRCKGPLPTDCHEGCAAGCTGPKHSDCLACLFHNS 240
D 202 SRCWGESSEDCQSLTRTYCAGGACRCKGPLPTDCHEGCAAGCTGPKHSDCLACLFHNS 261
QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289
D 262 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 14
AAB21203
ID AAB21203 standard; protein: 919 AA.
AC AAB21203;
DT 12-JAN-2001 (first entry)
XX Human HER-2/neu fusion protein.
DE
XX Human: HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer; fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX WO200044899-A1.
XX 03-AUG-2000.
PD 28-JAN-2000; 2000WO-US02164.
PF 29-JAN-1999; 99US-0117976.
XX
XX (CORI-) CORIXA CORP.
PA (SMIK) SMITHKLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX WPI: 2000-505976/45.
DR
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Claim 2; Fig 12; 128pp; English.
XX
CC The present sequence is a fusion protein comprising the extracellular
CC domain and the phosphorylation domain of the human HER-2/neu protein.
CC HER-2/neu is a member of the tyrosine kinase family of receptor-like
CC glycoproteins and shows homology to the epidermal growth factor receptor
CC (EGFR). It probably plays a part in cell growth and/or differentiation.
CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used
CC to treat or prevent cancer by eliciting or enhancing an immune response
CC to the HER-2/neu protein. They may be used to treat malignancies such as
CC breast, ovarian, colon, lung and prostate cancers, and may be used as an
CC antigen to vaccinate against these neoplasias.
XX
SQ Sequence 919 AA;

Query Match 100.0%; Score 1587; DB 21; Length 919;
Best Local Similarity 100.0%; Pred. No. 3,5e-131;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGYCTGDMKRLRPASPEETHLDMRLHLYGCGVVOGNGLELTYPPTNASLSFLDIOEVQ 60
D 22 STGYCTGDMKRLRPASPEETHLDMRLHLYGCGVVOGNGLELTYPPTNASLSFLDIOEVQ 81
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPTGASPGGLRELQ 120
D 82 GYVLIANQVQVPLQRLRIYRGTOLEFEDNYALAVLDGDBLNTTPTGASPGGLRELQ 141
QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRACHPSPCKG 180
D 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRACHPSPCKG 201
QY 181 SRCWGESSEDCQSLTRTYCAGGACRCKGPLPTDCHEGCAAGCTGPKHSDCLACLFHNS 240
D 202 SRCWGESSEDCQSLTRTYCAGGACRCKGPLPTDCHEGCAAGCTGPKHSDCLACLFHNS 261
QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289
D 262 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 15
AAM51148
ID AAM51148 standard; protein: 919 AA.
AC AAM51148;
DT 17-JUN-2002 (first entry)
XX
XX
DE Her-2/neu extracellular domain-phosphorylation domain fusion.
DE
KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
KW tyrosine kinase; receptor; c-erbB2; gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..653
FT /note= "extracellular domain"
FT 654..919
FT /note= "phosphorylation domain"
XX
PN WO200212341-A2.
PD 14-FEB-2002.
XX
PE 03-AUG-2001; 2001WO-US24283.
XX
PR 03-AUG-2000; 2000US-0632507.
XX
PA (CORI-) CORIXA CORP.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cheever MA, Gheysen D;
XX WPI: 2002-241743/29.
DR
XX Her-2/neu fusion protein for treating or preventing cancer by eliciting
PT or enhancing an immune response to the protein, has Her-2/neu
PT extracellular domain fused to Her-2/neu intracellular or
PT phosphorylation domain -
XX
PS Claim 2; Fig 12; 141pp; English.
XX
CC The present sequence is that of a fusion protein between the
CC extracellular domain and phosphorylation domain of human Her-2/neu
CC (see AAM51143), an oncogenic self-protein and target for anti-cancer
CC vaccines. The fusion protein can be obtained by recombinant DNA

CC methods. Her-2/neu overexpression correlates with a poor prognosis
 CC in breast and ovarian cancers. The invention provides Her-2/neu
 CC fusion proteins, nucleic acids encoding them, viral vectors, and
 CC vaccines comprising the fusion proteins or nucleic acid molecules.
 CC In preferred fusion proteins, the extracellular domain of a
 CC Her-2/neu protein is fused to a Her-2/neu intracellular domain or
 CC phosphorylation domain (or its Deltapd fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal ex vivo with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.
 CC
 XX

Sequence 919 AA:

Query Match 100.0%; Score 1587; DB 23; Length 919;
 Best Local Similarity 100.0%; Pred. No. 3.5e-131;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGYCTGDMKRLPASPETHLDMRLHYGCGVVGNGLETLPTNASLFLDIOEYQ 60
 DB 22 STGYCTGDMKRLPASPETHLDMRLHYGCGVVGNGLETLPTNASLFLDIOEYQ 81
 QY 61 GYVLIAHNOYROYPLQRLIRIVRGSTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 120
 DB 82 GYVLIAHNOYROYPLQRLIRIVRGSTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLRELO 141
 QY 121 LRSLETLKGGVLIQRPOLCYODTILMKDIFKNNQLALTLDITNRSRACHPCSPMCKG 180
 DB 142 LRSLETLKGGVLIQRPOLCYODTILMKDIFKNNQLALTLDITNRSRACHPCSPMCKG 201
 QY 181 SRMGESSEDCOSLTRVCGAGCARGPLPTDCHGQCAAGCTGPKHSDCLAFHNHS 240
 DB 202 SRMGESSEDCOSLTRVCGAGCARGPLPTDCHGQCAAGCTGPKHSDCLAFHNHS 261
 QY 241 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTPACPYNYLSTDVGS 289
 DB 262 GICELHCPALVTYNTDFFESMPNPEGRTFGASCVTPACPYNYLSTDVGS 310

Search completed: April 28, 2003, 13:39:46
 Job time : 26.1973 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:38:09 ; Search time 10.6321 Seconds

(without alignments)
2613.108 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STQVCTGTGTMKLRPASPER.....FGASCVTACPYNTLSTDVGS 289

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	1255	1 A24571	protein-tyrosine k
2	1353.5	85.3	1260	1 TVRRTU	epidermal growth f
3	1322	83.3	1254	2 I48161	P-185 precursor
4	739	46.6	527	2 A42032	epidermal growth f
5	739	46.6	1223	1 TVCHLV	epidermal growth f
6	716.5	45.1	1210	1 GQHUE	epidermal growth f
7	716.5	45.1	1210	2 A53183	epidermal growth f
8	716.5	45.1	1308	2 A47253	epidermal growth f
9	708.5	44.6	2	A36325	epidermal growth f
10	686	43.2	1342	2 A36223	kinase-related tra
11	652	41.1	1339	2 JC4387	epidermal growth f
12	640.5	40.4	1166	1 S06142	protein-tyrosine k
13	536	33.8	843	2 A27131	epidermal growth f
14	396.5	25.0	1323	2 E88257	protein let-23 (lm
15	396.5	25.0	1374	2 S70712	protein-tyrosine k
16	390.5	24.6	1369	2 S70713	epidermal growth f
17	375.5	23.7	1330	1 GQFEF	epidermal growth f
18	341	21.5	366	2 D45558	epidermal growth f
19	341	21.5	1717	1 A45558	epidermal growth f
20	331	20.9	333	2 B45558	epidermal growth f
21	331	20.9	342	2 C45558	epidermal growth f
22	315.5	19.9	1363	2 T43220	insulin-like growth
23	273.5	17.2	1477	2 T18534	insulin receptor
24	255	16.1	2101	2 S57245	insulin receptor
25	255	16.1	2148	1 A56081	insulin receptor
26	246.5	15.5	1300	2 A36502	insulin-like growth
27	246.5	15.5	1607	2 T43212	insulin-like growth
28	244.5	15.4	1383	2 A36080	insulin receptor p
29	243.5	15.3	1372	2 A34157	insulin receptor p

30	243.5	15.3	1382	1 INHUR	insulin receptor p
31	232.5	14.7	1367	1 IGHUR1	insulin-like growth
32	232.5	14.7	1371	2 A33837	insulin-like growth
33	231.5	14.6	1268	2 B36502	insulin receptor
34	228.5	14.4	340	2 B47417	insulin receptor
35	224.5	14.1	183	2 JY0803	tyrosine kinase re
36	220.5	13.9	329	2 A48805	insulin-like growth
37	207	13.0	1350	2 T30346	insulin receptor
38	185	11.7	1846	2 T42047	insulin receptor h
39	132.5	8.3	1299	2 T43251	insulin receptor h
40	125	7.9	1548	2 S34583	serine proteinase
41	121	7.6	1680	2 A43434	furin (EC 3.4.21.7
42	109.5	6.9	91	2 I51179	IGF-I receptor
43	108.5	6.8	798	2 B28193	integrin beta-1* c
44	107	6.7	915	1 A48225	subtilisin-like pr
45	106.5	6.7	837	2 S43656	furin (EC 3.4.21.7

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e
C:Species: Homo sapiens (man)
C>Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth
A:Reference number: A24571; MUID:66118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <TAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A25491; MUID:66016729; PMID:2955967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Cousens, I.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:66070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU2>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALU', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Molecule type: DNA
A:Status: translated from GB/EMBL/DBJ
A:Residues: 832-909 <REX>
A:Cross-references: GB:U29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Glyvol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:9183983; PIDN:AAA58637.1; PID:9553332
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERBB2; NGL: NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-17q21.1
 A:Introns: 25/1; 75/3; 147/1; 883/3
 A>Note: The list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
 F:22-653/Domain: extracellular #status predicted <EXT>
 F:70-104/Domain: EGF receptor extracellular domain repeat <EEL>
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
 F:654-675/Domain: transmembrane #status predicted <TM>
 F:676-1255/Domain: transmembrane #status predicted <TM>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase homology motif
 F:686/Binding site: phosphate (Thr) (covalent) (by tyrosine kinase C) #status predicted
 F:1133/Active site: Lys #status predicted
 F:1133,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 100.0%; Score 1587; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 2,6e-109;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TQVCTGTDKMLRLPASPETHLDMLRLHYOGCQVVGNGLETTYLPYNASISFLQDIOEVG 60
 DB 22 TQVCTGTDKMLRLPASPETHLDMLRLHYOGCQVVGNGLETTYLPYNASISFLQDIOEVG 81
 OY 61 GYVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYTGASPGGLRELO 120
 DB 82 GYVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYTGASPGGLRELO 141
 OY 121 LRSLEILKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNSRACGSPCKG 180
 DB 142 LRSLEILKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNSRACGSPCKG 201
 OY 181 SRGWGSSSEDCOSLRTVACGACGACGKGLPTDCCHQCAAGCTGPKHSDCLALHFHNS 240
 DB 202 SRGWGSSSEDCOSLRTVACGACGACGKGLPTDCCHQCAAGCTGPKHSDCLALHFHNS 261
 OY 241 GICELHCPALVYNTDTFESMPNREGRYTFGASCVTACAPYNYLSTDVGS 289
 DB 262 GICELHCPALVYNTDTFESMPNREGRYTFGASCVTACAPYNYLSTDVGS 310

RESULT 2
 TYRNTU
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C:Accession: A24562; #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 R:Barbault, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,
 Carcinogenesis 12, 1975-1978, 1991
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no
 2-thiozoyl[formamide or N-methyl-N-nitrosourea.
 A:Reference number: A61204; MUID:92035293; PMID:1682063
 A:Accession: A61204
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 637-663; 'V', 665-702 <MAS>
 A>Note: authors translated the codon GCA for residue 25 as Val
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>
 F:658-680/Domain: transmembrane #status predicted <TM>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:11,191,263,535,576,634/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 85.3%; Score 1353.5; DB 1; Length 1260;
 Best Local Similarity 85.3%; Pred. No. 4,4e-92;
 Matches 247; Conservative 13; Mismatches 28; Indels 1; Gaps 1;

OY 2 TQVCTGTDKMLRLPASPETHLDMLRLHYOGCQVVGNGLETTYLPYNASISFLQDIOEVG 61
 DB 26 TQVCTGTDKMLRLPASPETHLDMLRLHYOGCQVVGNGLETTYLPYNASISFLQDIOEVG 85
 OY 62 YVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYTGASPGGLRELO 120
 DB 86 YVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYTGASPGGLRELO 145
 OY 121 LRSLEILKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNSRACGSPCKG 180
 DB 146 LRSLEILKGVLIQNRNPOLCYODTILMKDIFKNNOLATLIDTNSRACGSPCKG 205
 OY 181 SRGWGSSSEDCOSLRTVACGACGACGKGLPTDCCHQCAAGCTGPKHSDCLALHFHNS 240
 DB 206 SRGWGSSSEDCOSLRTVACGACGACGKGLPTDCCHQCAAGCTGPKHSDCLALHFHNS 265
 OY 241 GICELHCPALVYNTDTFESMPNREGRYTFGASCVTACAPYNYLSTDVGS 289
 DB 266 GICELHCPALVYNTDTFESMPNREGRYTFGASCVTACAPYNYLSTDVGS 314

RESULT 3
 I48161
 P-185 precursor - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: I48161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishika
 Gene 140, 251-255, 1994
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
 A:Reference number: I48161; MUID:94193007; PMID:7908275
 A:Accession: I48161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1254 <RES>
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 83.3%; Score 1322; DB 2; Length 1254;
 Best Local Similarity 83.3%; Pred. No. 9,1e-90;
 Matches 240; Conservative 16; Mismatches 32; Indels 0; Gaps 0;

OY 2 TQVCTGTDKMLRLPASPETHLDMLRLHYOGCQVVGNGLETTYLPYNASISFLQDIOEVG 61
 DB 23 TQVCTGTDKMLRLPASPETHLDMLRLHYOGCQVVGNGLETTYLPANATISFLQDIOEVG 82
 OY 62 YVLIHNOVRQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTPTVYTGASPGGLRELO 121

```

Db 83 YMLAHSGVRAHPLQRLRIVRGTOLEEDKVALAVLDNDPDLNDVTATGRTPEGLREQL 142
Oy 122 RSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSPACHPCSPMCKGS 181
Db 143 RSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSPACHPCSPMCKGS 202
Oy 182 RCMGESSDCCSLTRTVACAGCA-RCKGRLPTDCHEOCACGTGPKHSDCLACHFNHSG 241
Db 203 HCMGASPEDCOTLTCTINAPRAVAPARARLPTDCHEOCACGTGPKHSDCLACHFNHSG 262
Oy 242 ICELHCPALVTYNTDTFESMPNPEGRTFGASCVTACPYNTLSDVGS 289
Db 263 ICELHCPALVTYNTDTFESMPNPEGRTFGASCVTACPYNTLSDVGS 310

RESULT 4
A42032
epidermal growth factor receptor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
R:Flickinger, T.W.; Mahle, N.J.; Kung, H.J.
Mol. Cell. Biol. 12, 883-893, 1992
A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble,
A:Reference number: A42032; MUID:92123214; PMID:1732751
A:Accession: A42032
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-527 <FL>
A:Cross-references: GB:M7637; NID:9211737; PIDN:AAA48759.1; PID:9211738
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBI:76892; NCBI:76893)
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor

Query Match 46.6%; Score 739; DB 2; Length 527;
Best Local Similarity 48.1%; Pred. No. 3.9e-47;
Matches 139; Conservative 48; Mismatches 90; Indels 12; Gaps 5;

Oy 3 QVCTGTDMKRLPASPEHLDMLRHLVYOGCOVGNLELTILPTNASLSFLDIDIEVGY 62
Db 38 KVCQGTNNKRLQLGHEHEDFTSLQRYNNCEVLSNLEITVEHNRDLTFLKTOIEVAGY 97
Oy 63 VLIANOVROYPLQRLRIVRGTOLEEDKVALAVLDNDPDLNDVTATGRTPEGLREQL 122
Db 98 VLIANOVROYPLQRLRIVRGTOLEEDKVALAVLDNDPDLNDVTATGRTPEGLREQL 148
Oy 123 SLEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSPACHPCSPMCKGS 181
Db 149 RSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSPACHPCSPMCKGS 207
Oy 182 RCMGESSDCCSLTRTVACAGCA-RCKGRLPTDCHEOCACGTGPKHSDCLACHFNHSG 240
Db 208 HCMGASPEDCOTLTCTINAPRAVAPARARLPTDCHEOCACGTGPKHSDCLACHFNHSG 267
Oy 241 ICELHCPALVTYNTDTFESMPNPEGRTFGASCVTACPYNTLSDVGS 289
Db 268 ATCKDTCPLVLYNPTTYOMDVNPEKYSFGATCVRECPHNVVTDHGS 316

RESULT 5
TVCNLY
epidermal growth factor receptor precursor - chicken
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
C:Species: Gallus gallus (chicken)
C:Date: 28-Feb-1996 #sequence_revision 05-May-1995 #text_change 04-Feb-2000
C:Accession: A27720; A00643
R:Iax, I.; Johnson, A.; How, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Vehnstr
Mol. Cell. Biol. 8, 1970-1978, 1988
A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in mou
A:Reference number: A27720; MUID:88261272; PMID:3260329
A:Accession: A27720

```

```

A:Molecule type: mRNA
A:Residues: 1-1223 <LAX>
A:Cross-references: GB:M20386
R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Kottman, F.M.; Crittenden, L.B.; Raines
Cell 41, 719-726, 1985
A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and
A:Reference number: A00643; MUID:85228222; PMID:2988784
A:Accession: A00643
A:Molecule type: mRNA
A:Residues: 585-1223 <NII>
A:Cross-references: GB:M10066
C:Gene: erbB
C:Superfamily: epidermal growth factor receptor; protein kinase homology;
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact
specific protein kinase
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>
F:31-1223/Domain: extracellular #status predicted <EXT>
F:81-307/Domain: EGF receptor extracellular domain repeat <EE1>
F:397-610/Domain: EGF receptor extracellular domain repeat <EE2>
F:655-677/Domain: transmembrane #status predicted <TM>
F:678-1223/Domain: intracellular #status predicted <INT>
F:719-984/Domain: protein kinase homology <KIN>
F:727-735/Region: protein kinase ATP-binding motif
F:136,202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent)
F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted
F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicte
F:754/Active site: Lys #status predicted
F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #s

Query Match 46.6%; Score 739; DB 1; Length 1223;
Best Local Similarity 48.1%; Pred. No. 9e-47;
Matches 139; Conservative 48; Mismatches 90; Indels 12; Gaps 5;

Oy 3 QVCTGTDMKRLPASPEHLDMLRHLVYOGCOVGNLELTILPTNASLSFLDIDIEVGY 62
Db 35 KVCQGTNNKRLQLGHEHEDFTSLQRYNNCEVLSNLEITVEHNRDLTFLKTOIEVAGY 94
Oy 63 VLIANOVROYPLQRLRIVRGTOLEEDKVALAVLDNDPDLNDVTATGRTPEGLREQL 122
Db 95 VLIANOVROYPLQRLRIVRGTOLEEDKVALAVLDNDPDLNDVTATGRTPEGLREQL 145
Oy 123 SLEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSPACHPCSPMCKGS 181
Db 146 RSLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNQALTLIDTNRSPACHPCSPMCKGS 204
Oy 182 RCMGESSDCCSLTRTVACAGCA-RCKGRLPTDCHEOCACGTGPKHSDCLACHFNHSG 240
Db 205 HCMGASPEDCOTLTCTINAPRAVAPARARLPTDCHEOCACGTGPKHSDCLACHFNHSG 264
Oy 241 ICELHCPALVTYNTDTFESMPNPEGRTFGASCVTACPYNTLSDVGS 289
Db 265 ATCKDTCPLVLYNPTTYOMDVNPEKYSFGATCVRECPHNVVTDHGS 313

RESULT 6
GORUE
epidermal growth factor receptor precursor - human
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
C:Species: Homo sapiens (man)
C:Date: 15-Nov-1984 #sequence_revision 27-Nov-1985 #text_change 11-Jun-1999
C:Accession: A00641; A25772; S30024; A38672; A43615; A23062; A05281; A60143;
R:Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J
rg, P.H.
Nature 309, 418-425, 1984
A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression
A:Reference number: A00641; MUID:84219729; PMID:6328312
A:Accession: A00641
A:Molecule type: mRNA
A:Residues: 1-1210 <ULL>
A:Cross-references: EMBL:X00588; NID:931113; PIDN:CAA25240.1; PID:9757924
A:Note: the authors translated the codon AAG for residue 540 as Asn

```

R: Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
A:Title: Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene
A:Reference number: A25772; MUID:85270438; PMID:2991899
A:Accession: A25772
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-29 <ISH>
A:Cross-references: GB:MI1234; NID:q181981; PIDN:AAA52370.1; PID:g553272
R: Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M.
Oncogene Res. 1, 375-396, 1987
A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification of exons
A:Reference number: S30024; MUID:88217353; PMID:3329716
A:Accession: S30024
A:Molecule type: DNA
A:Residues: 1-29 <HA2>
A:Cross-references: EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID:g31119
R: Haley, J.D.; Waterfield, M.D.
J. Biol. Chem. 266, 1746-1753, 1991
A:Title: Contributory effects of de Novo transcription and premature transcript termination
A:Reference number: A38672; MUID:91107677; PMID:1988448
A:Accession: A38672
A:Molecule type: DNA
A:Residues: 1-29 <HNL>
A:Cross-references: GB:M38425; NID:q181977; PIDN:AAA63171.1; PID:g553271
A:Experimental source: carcinoma cell line A431-7
R: Xu, Y.; Ishii, S.; Clark, A.J.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.; Merlino, I.
Nature 309, 806-810, 1984
A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RNAs
A:Reference number: A00642; MUID:84245855; PMID:6330563
A:Accession: A00642
A:Molecule type: mRNA
A:Residues: 'RCAMRRA', 150-187, 'KSVIGAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-321, '798-799, 'TP', 802-811, 'R', 813-942 <XUY>
A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF receptor
R: Lin, C.R.; Chen, W.S.; Krulger, W.; Scolarsky, L.S.; Weber, W.; Evans, R.M.; Verna, I.
Science 224, 843-846, 1984
A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplification
A:Reference number: A43615; MUID:84196372; PMID:6326261
A:Accession: A43615
A:Molecule type: mRNA
A:Residues: 713-964 <LIN>
A:Experimental source: epidermoid carcinoma cell line A431
R: Stimm, F.A.; Gope, M.L.; Schultz, T.2.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
Biochem. Biophys. Res. Commun. 124, 125-132, 1984
A:Reference number: A23062; MUID:85046483; PMID:6093780
A:Accession: A23062
A:Molecule type: mRNA
A:Residues: 1028-1210 <SIM>
R: Weber, W.; Gyll, C.N.; Speiss, J.
Science 224, 294-297, 1984
A:Reference number: A05281; MUID:84172183; PMID:6324343
A:Accession: A05281
A:Molecule type: protein
A:Residues: 25-30 'S', 32-51, 454-467 <WEP>
R: Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.
J. Biol. Chem. 260, 5205-5208, 1985
A:Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor
A:Reference number: A60143; MUID:85182650; PMID:2985580
A:Accession: A60143
A:Molecule type: protein
A:Residues: 740-744, 'X', 746-747 <RUS>
R: Roczko, B.; Mosig, G.; Cohen, S.
Nature 309, 270-273, 1984
A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide anion
A:Reference number: A38023; MUID:84191554; PMID:6325948
A:Accession: A38023
A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
R: Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.
Cell 59, 33-43, 1989
A:Title: Functional independence of the epidermal growth factor receptor from a domain A
A:Reference number: A3331; MUID:90003233; PMID:2790960
A:Accession: A3331
A:Contents: annotation, internalization signal

C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor
C:Genetics:
A:Gene: GDB:EGFR
A:Cross-references: GDB:120610; OMIM:131550
A:Map position: 7p12.3-7p12.1
A:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos
F:1-24/Domain: signal sequence #status predicted <SIG>
F:1-25-1210/Product: EGF receptor #status predicted <MAT>
F:25-645/Domain: extracellular #status predicted <EXT>
F:390-600/Domain: EGF receptor extracellular domain repeat <REP>
F:646-668/Domain: transmembrane #status predicted <TM>
F:669-1210/Domain: intracellular #status predicted <INT>
F:710-975/Domain: protein kinase homology <KIN>
F:718-726/Region: protein kinase ATP-binding motif
F:999-1046/Region: coated-pit mediated internalization signal
F:1047-1210/Region: inhibitory
F:128-175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status pre
F:745/Active site: Lys #status experimental
Query Match 45.1% Score 716.5; DB 1; Length 1210;
Best Local Similarity 46.5%; Pred. No. 4,1e-45;
Matches 134; Conservative 44; Mismatches 99; Indels 11; Gaps 2;
OY 3 QVCTGDMKRLPASPETHLMLRHLXGCGVQGNLELTPNLSLFLQDIQEVGY 62
DB 29 KVCQGTNRKLTQGTGFEDHFLSLQRMFNCEVVLGNLEIYVORNDLSPLKTIQEVAGY 88
OY 63 VLLAHNOYRVPYRORLRIVAGTOLFEDNVALVLDNDPLNNTPTVYGAAGGLREQLR 122
DB 89 VLLALNVERIPLENQIDIGNNYENSVALAVLSND-----AKRTGLKELPMR 138
OY 123 SLTEILKAGVLIORNPOLCYODTILMKDIPHKNNQTLALTLIDNRSRACRPSMCKGR 182
DB 139 NIDELHGNARFENNRNLCVESIDWIDYSSPFLSMNSMDPQNHLSGCKCPSCNGS 198
OY 183 CGESSEDCOSLTRFYCAGCA-RCKGRLPTDCCHQCAAGCTGPKHSDCLAEHNSG 241
DB 199 CMGAGEENCCKIKTKICACQCSGRCKSPSDCHNCAAGCTGPRSDCTVCRKFEDEA 258
OY 242 ICELHCPALVTYNDPESMPNPEGRTFGASCVCATCPYVLTSDVGS 289
DB 259 TCKDTCPMLNPTTYQMDVNEBGRKTSFGATCYKCKPRNYVTDHGS 306
RESULT 7
A53183
epidermal growth factor receptor precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
R: Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.
Genes Dev. 8, 399-413, 1994
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
A:Reference number: A53183; MUID:94170986; PMID:8125255
A:Accession: A53183
A:Molecule type: mRNA
A:Residues: 1-1210 <LUE>
A:Cross-references: GB:U03425
R: Aviv, A.; Lax, I.; Ullrich, A.; Schlesinger, J.; Givol, D.; Morse, B.
Oncogene 6, 673-676, 1991
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding
A:Reference number: A43818; MUID:91232866; PMID:2030916
A:Accession: A43818
A:Molecule type: mRNA
A:Residues: 1-714 <AVI>
A:Cross-references: GB:X59698
R: Bisinger, D.P.; Serrero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942
A:Molecule type: mRNA

OY 182 RCWGESSEDCQSLTRVYACGCA-RCKGPIPTDCEHCQACAGTGPKHSDCLAFHNHS 240
 DB 198 SCWGRGRENOCOKLTKIKKICRGRSPSCCHNOQACGTGCPRESCLVCHFRPDE 257
 OY 241 GICELHCPALVYNTDFEESMPNDEGRYTFGASCVTACPNYLTSDVGS 289
 DB 258 ATCKDTCPLMLYNTPTTQMDVNPDEKYSFGATCKKCPKRYVYVTDHGS 306

RESULT 10

A36223
 C:Species: Homo sapiens (man)
 C:Date: 04-Oct-1991 #sequence, revision 13-Jan-1993 #text, change 17-Nov-2000
 C:Accession: A36223; 159164
 R:Kraus, M.H.; Issing, W.; Miki, T.; Popescu, N.C.; Aaronson, S.A.
 Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
 A:Title: Isolation and characterization of ERB3, a third member of the ERBB/epidermal
 A:Reference number: A36223; MUID:90083234; PMID:2687875
 A:Accession: A36223
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1342 <KRA>

A:Cross-references: GB:M29366
 R:Blowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro, G.J.
 Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
 A:Title: Molecular cloning and expression of another epidermal growth factor receptor-
 A:Reference number: 159164; MUID:90311312; PMID:2164210
 A:Accession: 159164
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-555, 'G', 561-957, 'F', 959-1063, 'G', 1065-1342 <RES>
 A:Cross-references: GB:M34309; NID:g183990; PIDN:AAA35979.1; PID:g306841
 C:Genetics:

A:Gene: GDB:ERB3; HER3
 A:Cross-references: GDB:119880; OMIM:190151
 A:Map position: 12q13-12q13
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 C:Keywords: ATP; phosphotransferase
 F:707-972/Domain: protein kinase ATP-binding motif
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 43.2%; Score 686; DB 2: Length 1342;
 Best Local Similarity 46.1%; Pred. No. 8e-43; Indels 14; Gaps 5;
 Matches 131; Conservative 44; Mismatches 95

OY 1 STQVCTGTDKRLRPASPTHLDMLRHXGCVQVGNLELTLPYNASTFLQDIOEQV 60
 DB 25 SQAVCEPTLNGLSVTDGADNOGYOTLYKLYRCEVGNLEIYVIGHNADSLFQWIREVT 84
 OY 61 GYVLIANQVROYPLORLIRVGTQLEFDNYALAVLDNGDPLNNTPTVYGASPGGLELQ 120
 DB 85 GYVLIANNEFSTPLPLRLRVRGTOYDGFALFVM-----LWNT-----NSSHALROLK 135
 OY 121 LRSITELKGVLIORPOLCYODTILMKDIFKNNQALTLTIDNRSRACHCSPCKG 180
 DB 136 LTQLTLELSSGVITKDKLCHMDTIDMDIVDRD---AEIYVKNDGRSCPCHEVCKG 192
 OY 181 SRCWGESSEDCQSLTRTVACAGC-ARCKGPLPTDCHEQACAGCTGPKHSDCLAFHNH 239
 DB 193 -RCWGPESSEDCQSLTRTVACAGC-ARCKGPLPTDCHEQACAGCTGPKHSDCLAFHNH 251
 OY 240 SGICELHCPALVYNTDFEESMPNDEGRYTFGASCVTACPNYLT 283
 DB 252 SGACVPRCPPLVYNTDFEESMPNDEGRYTFGASCVTACPNYLT 295

RESULT 11

JC4387
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 17-Jan-1996 #sequence, revision 19-Apr-1996 #text, change 13-Nov-1998

C:Accession: JC4387
 R:Heiliger, N.J.; Kim, H.H.; Greaves, C.H.; Sterke, S.L.; Kojand, J.G.
 Gene 165, 279-284, 1995
 A:Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protein
 A:Reference number: JC4387; MUID:96096535; PMID:8522190
 A:Accession: JC4387
 A:Molecule type: mRNA
 A:Residues: 1-1339 <HELI>

A:Cross-references: GB:U29339; NID:g915389; PID:g915390
 A:Experimental source: liver
 A:Note: The authors translated the codon AAC for residue 369 as Thr and GTT for resid
 C:Comment: This protein is a functional heterologous receptor that transduces signals to
 C:Genetics:

A:Gene: ErbB3
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1339/Product: epidermal growth factor homolog #status predicted <TM>
 F:640-659/Domain: transmembrane #status predicted <TM>
 F:705-970/Domain: protein kinase ATP-binding motif
 F:713-721/Region: protein kinase ATP-binding motif
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr)

Query Match 41.1%; Score 652; DB 2: Length 1339;
 Best Local Similarity 44.4%; Pred. No. 2.6e-40; Indels 14; Gaps 5;
 Matches 126; Conservative 42; Mismatches 102

OY 1 STQVCTGTDKRLRPASPTHLDMLRHXGCVQVGNLELTLPYNASTFLQDIOEQV 60
 DB 25 SQAVCEPTLNGLSVTDGADNOGYOTLYKLYRCEVGNLEIYVIGHNADSLFQWIREVT 84
 OY 61 GYVLIANQVROYPLORLIRVGTQLEFDNYALAVLDNGDPLNNTPTVYGASPGGLELQ 120
 DB 85 GYVLIANNEFSTPLPLRLRVRGTOYDGFALFVM-----LWNT-----NSSHALROLK 135
 OY 121 LRSITELKGVLIORPOLCYODTILMKDIFKNNQALTLTIDNRSRACHCSPCKG 180
 DB 136 LTQLTLELSSGVITKDKLCHMDTIDMDIVDRD---AEIYVKNDGRSCPCHEVCKG 192
 OY 181 SRCWGESSEDCQSLTRTVACAGC-ARCKGPLPTDCHEQACAGCTGPKHSDCLAFHNH 239
 DB 193 -RCWGPESSEDCQSLTRTVACAGC-ARCKGPLPTDCHEQACAGCTGPKHSDCLAFHNH 251
 OY 240 SGICELHCPALVYNTDFEESMPNDEGRYTFGASCVTACPNYLT 283
 DB 252 SGACVPRCPPLVYNTDFEESMPNDEGRYTFGASCVTACPNYLT 295

RESULT 12

S06142
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 18-Feb-2000
 A:Accession: S06142; S13809
 R:Wittbrodt, J.; Adam, D.; Malitschek, B.; Maeueller, W.; Raulf, F.; Telling, A.; Robe
 Nature 341, 415-421, 1999
 A:Title: Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu
 A:Reference number: S06142; MUID:90015140; PMID:2797166
 A:Accession: S06142
 A:Molecule type: DNA
 A:Residues: 1-1166 <WIT>

A:Cross-references: EMBL:X16891; NID:g65290; PIDN:CAA34770.1; PID:g65291
 R:Adam, D.; Maeueller, W.; Scharf, M.
 Oncogene 6, 73-80, 1991
 A:Title: Transcriptional activation of the melanoma inducing Xmrk oncogene in Xiphoph
 A:Reference number: S13807; MUID:91125882; PMID:1846957
 A:Accession: S13809
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 821-1025, 'N', 1027-1098, 'A', 1100-1166 <ADA>
 A:Cross-references: EMBL:X56319; NID:g65284; PIDN:CAA39763.1; PID:g65285
 C:Genetics:

A:Gene: Xmrk
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1339/Product: epidermal growth factor homolog #status predicted <TM>
 F:640-659/Domain: transmembrane #status predicted <TM>
 F:705-970/Domain: protein kinase ATP-binding motif
 F:713-721/Region: protein kinase ATP-binding motif
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr)

Query Match 41.1%; Score 652; DB 2: Length 1339;
 Best Local Similarity 44.4%; Pred. No. 2.6e-40; Indels 14; Gaps 5;
 Matches 126; Conservative 42; Mismatches 102


```

A:Experimental source: strain N2
R:Koga, M.
submitted to the EMBL Data Library, July 1995
A:Reference number: S73101
A:Accession: S73101
A:Molecule type: DNA
A:Residues: 1-50, 'G', 52-1374 <KOG>
A:Cross-references: EMBL: D63426; NID: g1407562; PIDN: BAA09729.1; PID: g1407563
A:Experimental source: strain N2
R:Arizumi, R.V.; Koga, M.; Mendel, J.E.; Ohshima, Y.; Sternberg, P.W.
Nature 348, 693-699, 1990
A:Title: The let-23 gene necessary for Caenorhabditis elegans vulval induction encodes a
A:Reference number: S13422; MUID: g1080919; PMID: 1979659
A:Accession: S13422
A:Molecule type: mRNA
A:Residues: 52-1374 <ARO>
R:Thomas, K.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z20404
A:Accession: Z27662
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 52-1374 <WIL>
A:Cross-references: EMBL: Z70038; PIDN: CA93882.1; GSPDB: GNO0020; CESP: ZK1067.1
C:Genetic source: clone ZK1067
A:Gene: let-23; CESP: ZK1067.1
A:Map position: 2
A:Introns: 44/1; 51/1; 90/1; 112/3; 165/3; 217/1; 290/1; 379/1; 418/1; 448/2; 565/1; 606/1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homologs
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein kinase
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-1374/Product: protein-tyrosine kinase let-23 #status predicted <MAT>
F:934-1199/Domain: protein kinase homology <KIN>
F:942-950/Region: protein kinase ATP-binding motif

Query Match 25.0%; Score 396.5; DB 2; Length 1374;
Best Local Similarity 28.8%; Pred. No. 1.9e-21;
Matches 89; Conservative 54; Mismatches 113; Indels 53; Gaps 9;

OY 4 VCTGDMKRLRLPASPETHLMLRLHYGCGQVQVQGNLETLVLEPTN----- 47
DB 90 LCSGTGNGISIRYGTGNI-LEDELTMYRCGRVYGNLEITWEANIKKMRSTNSTVDPK 148
OY 48 -----ASLSFLODDOEVOGYLLAHNOYRQVPLRLRLIVRTQLEFENYALAVLDNGDP 101
DB 149 NEDSLPKSINFDNLEEIRGSLILYRANIKISFRLNVIYGVDFEHDN-ALYIHKNDK- 206
OY 102 LNNPTVPYTGASPGGLRLEQLRLTEILKGVLIQRPOLCY-ODTILKKDIFHKNNQAL 160
DB 207 -----VHEVYMRLEIRYIRNGSVTIQDNPKMCIYGDIDMKELLIDPD--YQ 250
OY 161 TLIDNRSRACH-----PCSPMCKGSRGCGESSEDCOSILRTVYACAGCARC--KGPL 210
DB 251 KVEETNSHQHCYONKSMACKHESC-NDCKWNSGDDNDCQRYRVSVCPRKSCQCFYNSGTS 309
OY 211 PFDCHBQCAAGCTGPRKSDCLAHFNHSGICELCALYTYNTDTESAPNPGRTTF 270
DB 310 SYECCDSACLGCGTGHGPKNCIACSKYELDGICTEPCSRKIFHNKRLRLVFNPDGRYON 369
OY 271 GASCVTACP 279
DB 370 GHHCVKBCP 378

```

Search completed: April 28, 2003, 13:42:29
Job time : 17.6321 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:23 ; Search time 4.53595 Seconds
(without alignments)
1984.228 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182
Sequence: 1 GAGGMVHRRHSSTRSGG.....STFKGPTAENPEYLGLDVP 217

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1177	99.6	1255	1	ERRB2_HUMAN
2	966.5	81.8	1234	1	ERRB2_MESAU
3	954	80.7	1257	1	ERRB2_RAT
4	283	23.9	245	1	ERRB2_MOUSE
5	164	13.9	1210	1	EGFR_HUMAN
6	157	13.3	1308	1	ERRB4_HUMAN
7	152	12.9	1308	1	ERRB4_RAT
8	145.5	12.3	707	1	SEPO_HUMAN
9	138	11.7	634	1	ERBBB_ALV
10	135.5	11.5	331	1	PRP1_HUMAN
11	135.5	11.5	553	1	ODO2_MYCTU
12	135	11.4	620	1	EXTN_TOBAC
13	134	11.3	1210	1	EGFR_MOUSE
14	130.5	11.0	886	1	SM6B_MOUSE
15	129.5	11.0	251	1	PRP2_HUMAN
16	128.5	10.9	1321	1	IRS2_MOUSE
17	127.5	10.8	3149	1	TEGU_EBV
18	125.5	10.6	2805	1	MAPA_HUMAN
19	124.5	10.5	775	1	ICPO_HSV1
20	124	10.5	503	1	MAIP_HUMAN
21	124	10.5	1106	1	GLI1_HUMAN
22	122	10.3	497	1	MAS2_HUMAN
23	122	10.3	518	1	TPM4_DROME
24	122	10.3	656	1	DNAA_STRCO
25	121.5	10.3	938	1	ENB4_EBV
26	120.5	10.2	449	1	APC_BRANA
27	120	10.2	296	1	PRP3_MOUSE
28	120	10.2	426	1	EXLP_TOBAC
29	120	10.2	699	1	VGIG_HSV2
30	119.5	10.1	296	1	CCO1_CAEEL
31	119.5	10.1	1219	1	YW29_YEAST
32	119	10.1	279	1	Y091_NPVP
33	119	10.1	439	1	XP2_XENLA

34	119	10.1	960	1	FGD1_MOUSE	P52734 mus musculus
35	118.5	10.0	604	1	RCO1_NEUR	P78706 neurospora
36	118	10.0	261	1	PRP2_MOUSE	P05142 mus musculus
37	118	10.0	419	1	SLY1_MOUSE	O61473 mus musculus
38	118	10.0	2167	1	SHK1_RAT	O94446 rattus norv
39	117.5	9.9	276	1	PRPL_HUMAN	P10162 homo sapien
40	117	9.9	1324	1	IRS2_HUMAN	O94442 homo sapien
41	117	9.9	1664	1	SLP1_CLOTM	O06852 clostridium
42	116.5	9.9	561	1	ERS_HUMAN	O43281 homo sapien
43	116	9.8	1248	1	DIAL_HUMAN	O60610 homo sapien
44	116	9.8	1274	1	ENAM_MOUSE	O55196 mus musculus
45	115.5	9.8	560	1	ERS_MOUSE	O64355 mus musculus

ALIGNMENTS

RESULT 1	ERRB2_HUMAN	STANDARD:	PRT: 1255 AA.
ID	ERRB2_HUMAN		
AC	P04626:		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	13-AUG-1987 (Rel. 05, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)		
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell		
DE	surface receptor HER2) (MLN 19).		
CN	ERRB2 OR HER2 OR NGL OR NEU.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86118663; PubMed=3003577;		
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,		
RA	Saito T., Toyoshima K.;		
RT	"Similarity of protein encoded by the human c-erbB-2 gene to		
RT	epidermal growth factor receptor.";		
RL	Nature 319:230-234(1986).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=86016729; PubMed=2995974;		
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,		
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,		
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor		
RT	shares chromosomal location with neu oncogene.";		
RL	Science 230:1132-1139(1985).		
RN	[3]		
RP	SEQUENCE OF 737-1031 FROM N.A.		
RX	MEDLINE=86016729; PubMed=2995974;		
RA	Semba K., Kanata N., Toyoshima K., Yamamoto T.;		
RA	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the		
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a		
RT	human salivary gland adenocarcinoma.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).		
RN	[4]		
RP	VARIANTS VAL-654 AND VAL-655.		
RX	MEDLINE=93194196; PubMed=8095488;		
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;		
RA	"Characterization of a new allele of the human ERBB2 gene by allele-		
RT	specific competition hybridization.";		
RL	Genomics 15:426-429(1993).		
CC	- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,		
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A		
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-		
CC	ALPHA AND AMPHIREGULIN.		
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein		
CC	tyrosine phosphate.		
CC	- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS		
CC	(POTENTIAL).		
CC	- SUBCELLULAR LOCATION: Type I membrane protein.		

CC -1- P-TM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY
 CC OF 0.783; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M11767; AAA35808.1; -.
 DR EMBL: M11761; AAA35808.1; JOINED.
 DR EMBL: M11762; AAA35808.1; JOINED.
 DR EMBL: M11763; AAA35808.1; JOINED.
 DR EMBL: M11764; AAA35808.1; JOINED.
 DR EMBL: M11765; AAA35808.1; JOINED.
 DR EMBL: M11766; AAA35808.1; JOINED.
 DR EMBL: M11730; AA475493.1; -.
 DR EMBL: M12036; AAA35978.1; -.
 DR EMBL: X03363; CAA27060.1; -.
 DR PIR: A25491; A25491.
 DR PIR: A24571; A24571.
 DR HSSP: P11362; 1FGK.
 DR Genew: HGNC:3430; ERBB2.
 DR MIM: 164870; -.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002174; Furin-1ike.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF00757; Furin-1like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TIR; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;
 KW Transmembrane; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Polymorphism.
 KM SIGNAL. 1 21
 FT CHAIN 22 1255 POTENTIAL.
 FT DOMAIN 22 652 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT TRANSMEM 653 675 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 676 1255 POTENTIAL.
 FT DOMAIN 720 987 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 726 987 PROTEIN KINASE.
 FT BINDING 753 734 ATP (BY SIMILARITY).
 FT ACT_SITE 845 845 ATP (BY SIMILARITY).
 FT DISULFID 199 204 BY SIMILARITY.
 FT DISULFID 220 227 BY SIMILARITY.
 FT DISULFID 224 235 BY SIMILARITY.
 FT DISULFID 236 244 BY SIMILARITY.
 FT DISULFID 240 252 BY SIMILARITY.
 FT DISULFID 255 264 BY SIMILARITY.
 FT DISULFID 268 295 BY SIMILARITY.
 FT DISULFID 299 311 BY SIMILARITY.
 FT DISULFID 315 331 BY SIMILARITY.
 FT DISULFID 334 338 BY SIMILARITY.
 FT DISULFID 351 320 BY SIMILARITY.
 FT DISULFID 515 528 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 544 560 BY SIMILARITY.

FT DISULFID 563 576 BY SIMILARITY.
 FT DISULFID 567 584 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 600 623 BY SIMILARITY.
 FT DISULFID 626 634 BY SIMILARITY.
 FT DISULFID 630 642 BY SIMILARITY.
 FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1248 1248 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 654 654 I -> V.
 FT VARIANT 654 654 /FTID-VAR_004077.
 FT VARIANT 655 655 I -> V.
 FT VARIANT 655 655 /FTID-VAR_004078.
 FT CONFLICT 1170 1170 P -> A (IN REF. 2).
 FT SEQUENCE 1255 AA; 137909 MM; 39E9DEDA04DC9662 CRC64;
 SO
 Query Match 99.6%; Score 1177; DB 1; Length 1255;
 Best Local Similarity 99.5%; Pred. No. 3.8e-64;
 Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAGGMVHRRSSSTRSGGDLTLGLPSESEAPRSPPLASSEGAGSDVPDGLGMAAG 60
 DB 1038 GAGGMVHRRSSSTRSGGDLTLGLPSESEAPRSPPLASSEGAGSDVPDGLGMAAG 1097
 QY 61 LQSLPTHDPSPLQRYSDPTVPPLPSETDGYVAPLTCSPQDEYVNPQVRRQPPSPRGPL 120
 DB 1098 LQSLPTHDPSPLQRYSDPTVPPLPSETDGYVAPLTCSPQDEYVNPQVRRQPPSPRGPL 1157
 QY 121 PAARPGATLEERAKTTLSPKNGVYKDVAFAGAVENREYLTPOGGAPOHPAPSPAF 180
 DB 1158 PAARPGATLEERKTLSPGKNGVYKDVAFAGAVENREYLTPOGGAPOHPAPSPAF 1217
 QY 181 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 217
 DB 1218 DNLVYWDQDPPERGAPSTFKGTPTAENPEYLGLDVP 1254
 RESULT 2
 ERB2_MESAU STANDARD; PRT; 1254 AA.
 ID ERB2_MESAU
 AC 060553;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (neu proto-oncogene) (C-erbB-2).
 GN ERBB2 OR NEU.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 OC Mesocricetus.
 OC NCBI_Taxid-10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Nerve;
 RX MEDLINE-94193007; PubMed-7908275;
 RA Nakamura T., Ushijima T., Ishizaka Y., Naga M., Arai M.,
 RA Yamazaki Y., Ishikawa T.;
 RT Cloning and activation of the Syrian hamster neu proto-oncogene.";
 RL Gene 140:251-255(1994).
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS

CC (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16295; BAA03801.1; -.
 DR HSSP: P11362; ITCG.
 DR InterPro: IPR000494; EGFRL_domain.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; Pkinase_1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR PRODOM: PD000001; Euk_Pkinase; 1.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 653 675 POTENTIAL.
 FT DOMAIN 676 1254 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 158 368 CYS-RICH.
 FT DOMAIN 472 644 CYS-RICH.
 FT DOMAIN 720 987 PROTEIN KINASE.
 FT NP_BIND 726 734 ATP (BY SIMILARITY).
 FT BINDING 753 753 ATP (BY SIMILARITY).
 FT ACT_SITE 845 845 BY SIMILARITY.
 FT DISULFID 195 204 BY SIMILARITY.
 FT DISULFID 199 212 BY SIMILARITY.
 FT DISULFID 236 244 BY SIMILARITY.
 FT DISULFID 240 252 BY SIMILARITY.
 FT DISULFID 255 264 BY SIMILARITY.
 FT DISULFID 268 295 BY SIMILARITY.
 FT DISULFID 299 311 BY SIMILARITY.
 FT DISULFID 315 331 BY SIMILARITY.
 FT DISULFID 334 338 BY SIMILARITY.
 FT DISULFID 511 520 BY SIMILARITY.
 FT DISULFID 515 528 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 544 560 BY SIMILARITY.
 FT DISULFID 563 576 BY SIMILARITY.
 FT DISULFID 567 584 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 600 623 BY SIMILARITY.
 FT DISULFID 626 634 BY SIMILARITY.
 FT DISULFID 630 642 BY SIMILARITY.
 FT MOD_RES 1139 1139 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).
 FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1254 AA; 138252 MW; 974C3791C21F2B81 CRC64;
 Query Match 81.8%; Score 966.5; DB 1; Length 1254;
 Best Local Similarity 82.0%; Pred. No. 1.8e-51;
 Matches 178; Conservative 13; Mismatches 25; Indels 1; Gaps 1;
 QY 1 GAGGVNHRHSSSTRSGGDLTGLPSEBAPSPPLASBEGASDVDFDGLGAAKG 60
 DB 1038 GAGSTAHRRHSSSTRSGGDLTGLPSEBAPSPPLASBEGASDVDFEGLGAAKG 1097
 QY 61 LQSLFTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNDVVRPQPPSPREGPL 120
 DB 1098 POSISPRDLSPQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNDVVRPQPPSPREGPL 1157
 QY 121 PAAPAGATLEBAKTLSTGKGVKADVAEGAVENPEYLPQGCAGAPQPPAPFSPAF 180
 DB 1158 PPVRPAGATLEBAKTLSTGKGVKADVAEGAVENPEYLPQGCAGAPQPPAPFSPAF 1216
 QY 181 DNLVYWDQDPPERGAPSTFGTPTAENPEYLGLDVP 217
 DB 1217 DNLVYWDQDPPERGAPSTFGTPTAENPEYLGLDVP 1253
 RESULT 3
 ERBB2_RAT
 ID ERBB2_RAT STANDARD: PRT: 1257 AA.
 AC P06494;
 DT 01-JAN-1998 (Rel. 06, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbb-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (epidermal growth factor receptor-related protein).
 DE ERBB2 OR NEU.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=86118662; PubMed=3945311;
 RA Bargmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related protein."
 RL Nature 319:226-230(1986).
 RN [2]
 RP SEQUENCE OF 852-905 FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=91222560; PubMed=2025425;
 RA Lal C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system."
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668
 RX MEDLINE=92155181; PubMed=1346763;
 RA Gullik W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D., Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-oncogenic and oncogenic forms of the neu protein."
 RL EMBO J. 11:43-48(1992).
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX. ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-ALPHA AND AMPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS. THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC PERSISTS IN THE EPITHELIUM AT BOTH THE IMPLANTATION AND
 CC INTERIMPLANTATION SITES IN ADDITION TO MODEST LEVELS IN THE
 CC SECONDARY DECIDUAL ZONE. ON DAYS 7 AND 8, ACCUMULATION IS ALSO
 CC PROMINENT IN THE TROPHOBLASTIC GIANT CELLS.
 CC -1- PPM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U71126; AAB17380.1; -
 CC EMBL: L47239; AAB93532.1; -
 CC HSSP: P11362; IFGK.
 CC MGD: MGI:95410; ErbB2.
 CC InterPro: IPR000719; Euk_Pkinase.
 CC InterPro: IPR004040; STY_Pkinase.
 CC InterPro: IPR001245; TYR_Pkinase.
 CC PRINTS: PR00109; TYRKINASE.
 CC PRODOM: PD000001; Euk_Pkinase; 1.
 CC SMART: SM00221; STYK; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; PARTIAL.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 CC Transmembrane; Glycoprotein; Multigene family; Receptor; Transferase;
 CC Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 CC NON_TER 1 1
 CC FT DOMAIN <1 >149 PROTEIN_KINASE.
 CC FT ACT_SITE 61 61 BY SIMILARITY.
 CC FT NON_CONS 149 150
 CC FT NON_TER 245 245
 CC SQ SEQUENCE 245 AA; 26927 MW; 0F763F0363DEFF1C CRC64;
 CC
 CC Query Match 23.9%; Score 283; DB 1; Length 245;
 CC Best Local Similarity 79.4%; Pred. No. 6.2e-11;
 CC Matches 54; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
 CC
 CC QY 1 GAGGVNHRHRSSTRSGGDLTGLEPSEEPAPRSPAPSEAGSDVFDGLGKAAG 60
 CC Db 177 GTGSAHRHRHRSSTRSGGDLTGLEPSEEPAPRSPAPSEAGSDVFDGLGKAAG 236
 CC QY 61 LQSLPETHD 68
 CC Db 237 LQSLPETHD 244
 CC
 CC RESULT 5
 CC EGF_R_HUMAN STANDARD; PRT; 1210 AA.
 CC AC P00533; P06268; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q92795; O00732;
 CC AC O00683; Q9BZS2; Q9BZC9; Q9H2C9; Q9H3C9;
 CC DT 21-JUL-1986 (Ref. 01, Created)
 CC DT 01-NOV-1997 (Ref. 35, Last sequence update)
 CC DT 15-JUN-2002 (Ref. 41, Last annotation update)
 CC DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor
 CC protein-tyrosine kinase ErbB-1).
 CC GN EGF_R OR ERBB1.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC OX NCI-ThxID=9606;
 CC RN NCBI
 CC RP SEQUENCE FROM N.A. (ISOFORM 1).
 CC RX MEDLINE-84219729; PubMed-6328312;
 CC RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,
 CC Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,
 CC Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;
 CC *Human epidermal growth factor receptor cDNA sequence and aberrant

RT expression of the amplified gene in A431 epidermoid carcinoma cells.";
 RL Nature 309:418-425(1984).
 RL [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE-95382957; PubMed-7654368;
 RA Ilekis J.V., Stark B.C., Scoccia B.;
 RT "Possible role of variant RNA transcripts in the regulation of
 RT epidermal growth factor receptor expression in human placenta.";
 RL Mol. Reprod. Dev. 41:149-156(1995).
 RL [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE-97076686; PubMed-9918811;
 RA Reiter J.L., Malhale N.J.;
 RT "A 1.8 kb alternative transcript from the human epidermal growth
 RT factor receptor gene encodes a truncated form of the receptor.";
 RL Nucleic Acids Res. 24:4050-4056(1996).
 RL [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE-Placenta;
 RX MEDLINE-97256547; PubMed-9103388;
 RA Ilekis J.V., Garfili J., Niederberger C., Scoccia B.;
 RT "Expression of a truncated epidermal growth factor receptor-like
 RT protein (TEGFR) in ovarian cancer.";
 RL Gynecol. Oncol. 65:36-41(1997).
 RL [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
 RC TISSUE-Placenta;
 RX MEDLINE-21100872; PubMed-11161793;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
 RA Scheel Sinclair C., Peersall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,
 RA Malhale N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative EGF receptor transcripts encoding truncated receptor
 RT isoforms.";
 RL Genomics 71:1-20(2001).
 RL [6]
 RP SEQUENCE OF 575-687 FROM N.A.
 RA Reiter J.L., Threadgill D.W., Danielsen A.J., Scheel C.M.,
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
 RA Malhale N.J.;
 RT "Human and mouse alternative EGF receptor transcripts encoding only the
 RT extracellular domain of the receptor.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RL [7]
 RP SEQUENCE OF 713-924 FROM N.A.
 RX MEDLINE-84196372; PubMed-6326261;
 RA Lin C.R., Chen W.S., Kruller W., Stolarsky L.S., Weber M.,
 RA Evans R.M., Verma I.M., Gill G.N., Rosenfeld M.G.;
 RT "Expression cloning of human EGF receptor complementary DNA: gene
 RT amplification and three related messenger RNA products in A431
 RT cells.";
 RL Science 224:843-848(1984).
 RL [8]
 RP SEQUENCE OF 150-962 FROM N.A.
 RX MEDLINE-84245835; PubMed-6330563;
 RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,
 RA Roe B.A., Merlino G.T., Pastan I.;
 RT "Human epidermal growth factor receptor cDNA is homologous to a
 RT variety of RNAs overproduced in A431 carcinoma cells.";
 RL Nature 309:806-810(1984).
 RL [9]
 RP SEQUENCE OF 1028-1210 FROM N.A.
 RX MEDLINE-85046483; PubMed-6093780;
 RA Simmen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,
 RA O'Malley B.W.;
 RT "Isolation of an evolutionarily conserved epidermal growth factor
 RT receptor cDNA from human A431 carcinoma cells.";
 RL Biochem. Biophys. Res. Commun. 124:125-132(1984).
 RL [10]
 RP SEQUENCE OF 1-29 FROM N.A.

RX MEDLINE-88217333; PubMed-3329716;
 RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,
 RA Waterfield M.D.;
 RT "The human EGF receptor gene: structure of the 110 kb locus and
 RT identification of sequences regulating its transcription.";
 RL Oncogene Res. 1:375-396(1987).
 RN [111]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE-91107677; PubMed-1988448;
 RA Haley J.D., Waterfield M.D.;
 RT "Contributory effects of de novo transcription and premature
 RT transcription termination in the regulation of human epidermal growth
 RT factor receptor proto-oncogene RNA synthesis.";
 RL J. Biol. Chem. 266:1746-1753(1991).
 RN [112]
 RP SEQUENCE OF 1-29 FROM N.A.
 RX MEDLINE-85270438; PubMed-2991899;
 RA Ishii S., Xu Y.H., Stratton R.H., Roe B.A., Merlino G.T., Pastan I.;
 RT "Characterization and sequence of the promoter region of the human
 RT epidermal growth factor receptor gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
 RN [113]
 RP SEQUENCE OF 540.
 RA Kohda D.;
 RL Submitted (SEP-1997) to the SWISS-PROT data bank.
 RN [114]
 RP RECEPTOR ACTIVITY.
 RX MEDLINE-84191554; PubMed-6325948;
 RA Mroczkowski B., Mosig G., Cohen S.;
 RT "ATP-stimulated interaction between epidermal growth factor receptor
 RT and supercoiled DNA";
 RL Nature 309:270-273(1984).
 RN [115]
 RP PHOSPHORYLATION.
 RX MEDLINE-89278137; PubMed-2543678;
 RA Margolis B.L., Lax I., Kris R., Dombalagian M., Honegger A.M.,
 RA Howk R., Glyol D., Ullrich A., Schlessinger J.;
 RT "All autophosphorylation sites of epidermal growth factor (EGF)
 RT receptor and HER2/neu are located in their carboxyl-terminal tails.
 RT Identification of a novel site in EGF receptor.";
 RL J. Biol. Chem. 264:10667-10671(1989).
 RN [116]
 RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND
 RP ASN-528.
 RX MEDLINE-96398132; PubMed-8962717;
 RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;
 RT "Analysis of the glycosylation patterns of the extracellular domain of
 RT the epidermal growth factor receptor expressed in Chinese hamster
 RT ovary fibroblasts";
 RL Growth Factors 13:121-132(1996).
 RN [117]
 RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND
 RP ASN-603.
 RX MEDLINE-20198209; PubMed-10731668;
 RA Sato C., Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;
 RT "Characterization of the N-oligosaccharides attached to the atypical
 RT Asn-X-Cys sequence of recombinant human epidermal growth factor
 RT receptor.";
 RL J. Biochem. 127:65-72(2000).
 RN [118]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RX MEDLINE-98225196; PubMed-9556602;
 RA Abe Y., Otake M., Inagaki F., Lax I., Schlessinger J., Kohda D.;
 RT "Disulfide bond structure of human epidermal growth factor receptor.";
 RL J. Biol. Chem. 273:11150-11157(1998).
 RN [119]
 RP REVIEW.
 RX MEDLINE-87297456; PubMed-3039909;
 RA Carpenter G.;
 RT "Receptors for epidermal growth factor and other polypeptide
 RT mitogens.";
 RL Annu. Rev. Biochem. 56:881-914(1987).
 CC -I- FUNCTION: Receptor for EGF, but also for other members of the EGF

CC family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding
 CC EGF-like growth factor, gp30 and vaccine virus growth factor. Is
 CC involved in the control of cell growth and differentiation.
 CC -I- FUNCTION: Isoform 2/truncated isoform may act as an antagonist.
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is
 CC secreted.
 CC -I- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/
 CC truncated isoform/TEGFR, 3/p110 and 4; are produced by
 CC alternative splicing.
 CC -I- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also
 CC expressed in ovarian cancers.
 CC -I- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X00588; CAA25240.1; -
 CC DR EMBL: U95089; AAB53063.1; -
 CC DR EMBL: U48723; AAC50802.1; -
 CC DR EMBL: U48723; AAC50804.1; -
 CC DR EMBL: U48724; AAC50796.1; -
 CC DR EMBL: U48725; AAC50797.1; -
 CC DR EMBL: U48726; AAC50798.1; -
 CC
 CC Query Match 13.9%; Score 164; DB 1; Length 1210;
 CC Best Local Similarity 30.5%; Pred. No. 0.004%; Gaps 11;
 CC Matches 62; Conservative 24; Mismatches 69; Indels 48; Gaps 11;
 CC
 CC 29 SEEEAPRSLPAPSGAGSDVFDGDLGMAKGLQSLTPHDSPLQRYSEDPYLPSET- 87
 CC 1025 SSPSTSTPLSLSSASNS--NSTVACIDRNGQSCIKRDSFLQRRSSPTGLTSDSI 1082
 CC
 CC 88 -DGVADPLTCSPOPEYVNDVDRPQPPSPREGLPAPRPGATLIERAKTISPCKNGYVKD 146
 CC 1083 DDFLL-----PVPEYINQ--SVPRPAGSVQNPVYHNPPLNP-----APSRDPHYQD 1127
 CC
 CC 147 VFAPFGAVENPEYL-TPGGAAAPQHPHPAPFAFDLXYWDQ-----DP----- 190
 CC 1128 --PHSTAVGNPEYLTNQ-----PTCVNSTPSPAHMAQSGSHQISLDNDPYQQDF 1176
 CC
 CC QY 191 -PERGAPPTFKGTPTAENPEYL 212
 CC DB 1177 FPKFAKNGIFKKS-TAENAEYL 1198
 CC
 CC RESULT 6
 CC ERB4_HUMAN
 CC ID ERB4_HUMAN STANDARD; PRT; 1308 AA.
 CC AC Q15303;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)
 CC DE (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).
 CC GN ERB4 OR HER4
 CC OS Homo sapiens (human).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC RP SEQUENCE FROM N.A. (ISOFORM JM-A).
 CC RC TISSUE-Breast carcinoma;

RX MEDLINE-93189574; PubMed-8383326;
 RA Plozman G.D., Culouscog J.-M., Whitney G.S., Green J.M., Carlton G.W.,
 RA Roy L., Neuhauer M.G., Shoyab M.;
 RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the
 RT epidermal growth factor receptor family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).
 RC TISSUE-Fetal brain;
 RX MEDLINE-97476287; PubMed-9334263;
 RA Elenius K., Cortas G., Paul S., Choi C.-J., Rio C., Plozman G.D.,
 RA Klagsbrun M.;
 RT "A novel juxtamembrane domain isoform of HER4/erbB4. Isoform-specific
 RT tissue distribution and differential processing in response to
 RT phorbol ester".
 RL J. Biol. Chem. 272:26761-26768(1997).
 CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
 CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND
 CC NTAK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.
 CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
 CC RECEPTORS (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: JM-A (SHOWN HERE) AND JM-B;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER
 CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED
 CC BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND
 CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN
 CC CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,
 CC KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,
 CC PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,
 CC LUNG, SALIVARY GLAND, AND PANCREAS.
 CC -1- PFM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -1- THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L07868; AAB59446.1; -
 DR HSSP: P11362; IEGK.
 DR Genew: HGNC:3432; ERBB4.
 DR MIM: 600543; -
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR000719; Euk_Kinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_Kinase.
 DR InterPro: IPR004019; YLP_molclif.
 DR Pfam: PF00069; pkinase.1.
 DR Pfam: PF00757; Furin-like.1.
 DR Pfam: PF01030; Recep_L.domain.2.
 DR Pfam: PF02757; YLP.2.
 DR PRODOM: PD000001; Euk_pkinase.1.
 DR SMART: SM00261; FU; 4.
 DR SMART: SM00219; TyrcK.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Alternative splicing.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 1308 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.
 FT DOMAIN 26 651 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 652 675 POTENTIAL.

FT DOMAIN 676 1308 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 186 334 CYS-RICH.
 FT DOMAIN 496 633 CYS-RICH.
 FT DOMAIN 718 985 PROTEIN KINASE.
 FT NP_BIND 724 732 ATP (BY SIMILARITY).
 FT BINDING 751 751 ATP (BY SIMILARITY).
 FT ACT_SITE 843 843 BY SIMILARITY.
 FT DISULFID 189 197 BY SIMILARITY.
 FT DISULFID 193 205 BY SIMILARITY.
 FT DISULFID 213 221 BY SIMILARITY.
 FT DISULFID 217 229 BY SIMILARITY.
 FT DISULFID 230 238 BY SIMILARITY.
 FT DISULFID 234 246 BY SIMILARITY.
 FT DISULFID 249 258 BY SIMILARITY.
 FT DISULFID 262 289 BY SIMILARITY.
 FT DISULFID 293 304 BY SIMILARITY.
 FT DISULFID 308 323 BY SIMILARITY.
 FT DISULFID 326 330 BY SIMILARITY.
 FT DISULFID 503 512 BY SIMILARITY.
 FT DISULFID 507 520 BY SIMILARITY.
 FT DISULFID 523 532 BY SIMILARITY.
 FT DISULFID 536 552 BY SIMILARITY.
 FT DISULFID 555 569 BY SIMILARITY.
 FT DISULFID 559 577 BY SIMILARITY.
 FT DISULFID 580 589 BY SIMILARITY.
 FT DISULFID 593 614 BY SIMILARITY.
 FT DISULFID 617 625 BY SIMILARITY.
 FT DISULFID 621 633 BY SIMILARITY.
 FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 626 648 NGPSTHDCIYPTMGHSTLPDHA -> IGSSIEDICIGLMD
 FT VARSPLIC (IN ISOFORM JM-B).
 FT FT
 SO SEQUENCE 1308 AA; 146807 MW; 5E4AE80985D88761 CRC64;
 Query Match 13.38; Score 157; DB 1; Length 1308;
 Best Local Similarity 24.28; Pred. No. 0.012;
 Matches 62; Conservative 23; Mismatches 75; Indels 96; Gaps 10;
 QY 16 RSGGDLTLGLE-PSEEAAPRSLAP-SEGAGSDVDDGLMGAKKQSLPTHPSPLO 73
 DB 1067 RDGGAAGQGVSVYPRAPTSTIPAPVAQATAEFTDSCCGNTRAKVAPRVQEDSSIQ 1126
 QY 74 RYSDPTVPLPS-----ETDGYAAPLTGSPQPEYVQNPQVPRQPPSPRGGLPAABA 126
 DB 1127 RYSDPTVFAFAPERSPRGELDEGVWTMRDKPKQGYLNPVE-----ENPFVSR-- 1175
 QY 127 GATLERAKTLSPGKGVKDYFAFGGAVENPEYLTPOGGAAPQPIPPA----- 175
 DB 1176 -----KNGDQ-----ALDNPENHNASNG-----PPRAEDYVNEPLYL 1209
 QY 176 -----FSPADNLYWMDPPERGA--PSTFNGTPT----- 205
 DB 1210 NTFANTLCKAEYLNKNNIISMPEKAKKAPDNDPYMHSI--PPSTLDHPDYLCEYSTKYTK 1269
 QY 206 -----AENPEYL 212
 DB 1270 QNGRIRPIVAENPEYL 1285
 RESULT 7

ERB4_RAT
ID ERB4_RAT STANDARD; PRT; 1308 AA.
AC Q62956; Q922N7;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
GN ERB4 OR TYRO-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=98221155; PubMed=9553078;
RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
RA Marchionni M.A., Kelly R.A.;
RT "Neuregulins promote survival and growth of cardiac myocytes.
RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT ventricular myocytes.";
RL J. Biol. Chem. 273:10261-10269(1998).
RN [2]
RP SEQUENCE OF 848-901 FROM N.A.
RC TISSUE=Sciatic nerve;
RX MEDLINE=91223560; PubMed=20253425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system.";
RL Neuron 6:691-704(1991).
RN [3]
RP SEQUENCE OF 1031-1198 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spinal cord;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Frohnert P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuregulins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration.";
RL J. Neurosci. 17:1642-1659(1997).
RN [4]
RP FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
RN 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND
RN NTA. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.
RN NOT ACTIVATED BY EGF, TGF- α , AND AMPHIREGULIN (BY SIMILARITY).
RN -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
RN tyrosine phosphate.
RN -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERB
RN RECEPTORS (POTENTIAL).
RN -1- SUBCELLULAR LOCATION: Type I membrane protein.
RN -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING
RN NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS
RN OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE
RN RETICULAR NUCLEUS OF THE THALAMUS, VERY LOW LEVELS IN KIDNEY, AND
RN HEART.
RN -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
RN RESIDUES (BY SIMILARITY).
RN -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RN or send an email to license@isb-sib.ch).
RN -----
RN EMBL: AF041838; AAD08899.1; -
RN EMBL: U52531; AAC53051.1; -
RN HSP: P11362; 1FCR.
RN InterPro: IPR000494; EGFR_L_domain.
RN InterPro: IPR000719; Euk_pkinase.
RN InterPro: IPR002174; Furin-like.
RN InterPro: IPR001245; Tyr_kinase.
RN InterPro: IPR004019; YLP_motif.
RN Pfam: PF00757; Furin-like; 1.
RN

DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; F0; 4.
DR SMART: SM00219; TYRKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;
KW Transmembrane; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT TRANSFERASE; 1
FT SIGNAL; 1
FT CHAIN; 1
FT DOMAIN; 26
FT DOMAIN; 26
FT TRANSMEM; 652
FT DOMAIN; 675
FT DOMAIN; 675
FT DOMAIN; 1308
FT DOMAIN; 1308
FT DOMAIN; 1308
FT DOMAIN; 1308
FT DOMAIN; 1308
FT NP_BIND; 724
FT BINDING; 724
FT ACT_SITE; 751
FT ACT_SITE; 751
FT DISULFID; 843
FT DISULFID; 843
FT DISULFID; 189
FT DISULFID; 193
FT DISULFID; 205
FT DISULFID; 213
FT DISULFID; 221
FT DISULFID; 229
FT DISULFID; 230
FT DISULFID; 234
FT DISULFID; 246
FT DISULFID; 249
FT DISULFID; 262
FT DISULFID; 289
FT DISULFID; 293
FT DISULFID; 304
FT DISULFID; 308
FT DISULFID; 323
FT DISULFID; 326
FT DISULFID; 330
FT DISULFID; 330
FT DISULFID; 503
FT DISULFID; 507
FT DISULFID; 520
FT DISULFID; 523
FT DISULFID; 536
FT DISULFID; 555
FT DISULFID; 559
FT DISULFID; 577
FT DISULFID; 580
FT DISULFID; 589
FT DISULFID; 593
FT DISULFID; 614
FT DISULFID; 617
FT DISULFID; 625
FT DISULFID; 621
FT MOD_RES; 1162
FT MOD_RES; 1162
FT MOD_RES; 1188
FT MOD_RES; 1258
FT MOD_RES; 1258
FT MOD_RES; 1284
FT CARBOHYD; 138
FT CARBOHYD; 138
FT CARBOHYD; 174
FT CARBOHYD; 174
FT CARBOHYD; 253
FT CARBOHYD; 253
FT CARBOHYD; 358
FT CARBOHYD; 358
FT CARBOHYD; 410
FT CARBOHYD; 410
FT CARBOHYD; 473
FT CARBOHYD; 473
FT CARBOHYD; 495
FT CARBOHYD; 495
FT CARBOHYD; 348
FT CARBOHYD; 348
FT CARBOHYD; 576
FT CARBOHYD; 576
FT CARBOHYD; 620
FT CARBOHYD; 620
FT CONFLICT; 1062
FT CONFLICT; 1062
FT CONFLICT; 1080
FT CONFLICT; 1080
SQ SEQUENCE 1308 AA; 146957 MW; D944B0996A08B41 CRC64;
Query Match 12.9%; Score 152; DB 1; Length 1308;
Best Local Similarity 24.2%; Pred. No. 0.024;
Matches 57; Conservative 23; Mismatches 60; Indels 96; Gaps 9;
OY 34 PRSPLAPSEAGSDVFDGDLGMAAGLGLPHRDSPLQRYSEDPTPLP-----E 86
DB 1089 PEARPA--QGATFEMPDSCNCTLRKPVVPHVQEDSSQRYADPTVFAERNPRAELD 1146
OY 87 TDGYVAPLTCSPQPEVYNQDPVROPSPREGCLPAARPGATLERAKTLSPGKNGVKD 146
DB 1147 EBGYMPMDHDKPKQETLNPVE-----ENPFVSR-----KNGDLO 1181

```

OY 147 VFAAGAVENPEYLTPOGGAAPQPPPPA-----175
DB 1182 -----ALDNEFYHSASSG-----PKAEDEYNEPLYLNTFTNALGNAEYMKNSLSLV 1229
OY 176 ---FSPAFDNLVYWDODPPERGA---PPSTFKGTP-----AENPEYL 212
DB 1230 PEKAKKAFNDNDYNNHSLPRSTLQHPDYLDQYSTKVFYKONGHRIPIVAENPEYL 1285

RESULT 8
SFPO_HUMAN
ID SFPO_HUMAN STANDARD; PRT; 707 AA.
AC P23246; P30808;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Splicing factor, proline- and glutamine-rich (polypyrimidine tract-
DE binding protein-associated splicing factor) (PTB-associated splicing
DE factor) (PSF) (DNA-binding p52/p100 complex, 100 kDa subunit).
OS SFPO OR PSF.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND ALTERNATIVE SPLICING.
RC TISSUE=Fetal brain;
RA MEDLINE=93194059; PubMed=8449401;
RT Patton J.G., Porro E.B., Galceran J., Tempst P., Nadal-Ginard B.;
RT "Cloning and characterization of PSF, a novel pre-mRNA splicing
RT factor."
RL Genes Dev. 7:393-406(1993).
RN [2]
RP SEQUENCE OF 312-707 FROM N.A.
RC TISSUE=Fetal skeletal muscle;
RA MEDLINE=90091812; PubMed=2480877;
RT Gower H.J., Moore S.E., Dickson G., Elsom V.L., Nayak R., Walsh F.S.;
RT "Cloning and characterization of a myoblast cell surface antigen
RT defined by 24 ID5 monoclonal antibody."
RL Development 105:723-731(1989).
RN [3]
RP SEQUENCE OF 48-68 AND 213-246.
RA MEDLINE=93176127; PubMed=8439294;
RT Zhang W.-W., Zhang L.-X., Busch R.K., Fairies J., Busch H.;
RT "Purification and characterization of a DNA-binding heterodimer of 52
RT and 100 kDa from HeLa cells."
RL Biochem. J. 290:267-272(1993).
CC -1- FUNCTION: ESSENTIAL PRE-MRNA SPLICING FACTOR REQUIRED EARLY IN
CC SPLICOSOME FORMATION. BINDS TO THE MAMMALIAN POLYPRIMIDINE
CC TRACTS. FORMS A COMPLEX WITH THE POLYPRIMIDINE TRACT-BINDING
CC PROTEIN (PTB). SEEMS TO ALSO BIND DNA.
CC -1- SUBUNIT: HETEROTETRAMER OF TWO 52 KDA AND TWO 100 KDA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; A LONG FORM (SHOWN
CC HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM)
CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE MYOBLAST CELL
CC SURFACE ANTIGEN 24.ID5 AND A POSSIBLE MEMBRANE-BOUND PROTEIN
CC ECTOKINASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70944; CA50283.1; -
DR EMBL: X16850; CA34747.1; -
DR PIR: A43557; A43557.
DR PIR: S29770; S29770.

```

```

DR HSSP; P11940; ICVJ.
DR SWISS-2DPAGE: P23246; HUMAN.
DR Genew: HGNC:10774; SFPO.
DR MIM: 605199; -.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS00360; RRM; 2.
DR PROSITE: PS00030; RRM_RNP_1; 1.
DR Nucleic protein; RNA-binding; DNA-binding; mRNA splicing; Repeat;
KW Alternative splicing.
FT DOMAIN 297 369 RNA-BINDING (RRM) 1.
FT DOMAIN 371 452 RNA-BINDING (RRM) 2.
FT DOMAIN 9 27 3 X 3 AA REPEATS OF R-G-G.
FT REPEAT 19 21 1.
FT REPEAT 25 27 2.
FT REPEAT 10 10 3.
FT DOMAIN 10 15 GLN/GLU/PRO-RICH.
FT DOMAIN 20 27 POLY-GLY.
FT DOMAIN 56 65 POLY-GLY.
FT DOMAIN 67 71 POLY-PRO.
FT DOMAIN 95 98 POLY-GLN.
FT DOMAIN 99 103 POLY-GLN.
FT DOMAIN 184 188 POLY-PRO.
FT DOMAIN 571 574 POLY-ARG.
FT DOMAIN 613 616 POLY-GLY.
FT DOMAIN 635 641 POLY-GLY.
FT VARSPLIC 663 707 POLY-GLY.
FT CONFLICT 243 243 RTERFGGAGPVGCGGCPGRCGCTPAGYGRGREG
FT FT PNKRPRF -> VRMIDVG (IN SHORT ISOFORM).
FT FT G -> R (IN REF. 3).
SQ SEQUENCE 707 AA; 76149 MW; 6D8D5EA35E235847 CRC64;

Query Match 12.3%; Score 145.5; DB 1; Length 707;
Best Local Similarity 26.6%; Pred. No. 0.033;
Matches 58; Conservative 13; Mismatches 82; Indels 65; Gaps 10;

OY 1 GAGGAVHRRHSSSTRSGGDLTGLPESEDEAPRSLAPSEGAGSDVFDGLGKAAG 60
DB 11 GGGGCFHRR-----GGGGRGGLNDFR-----SPPGAGLNNRPGMKRPGPOSS 54
OY 61 LQSLPTHDPSPLQRYSE-DPTVPLPSETDGYVAPLTCSPOPEYVNPQDVRQPPSPRE- 117
DB 55 GPKRPPIPPPPHQQQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQSS 107
OY 118 -----GPIPARPACATIERAKTILSPKNGVYKDVFAAGAVENPEYLTPOGGAAPQ 169
DB 108 KPVYAQGPAPGVGSAPASSAAPPATPTSGA-----PPGSGPGPT 150
OY 170 PHPPAFSPAFDNLVYWDODPPERGAPST--FKGTP 205
DB 151 PTPPPAVTSA-----PP--GAPPTPPSSGVPT 176

RESULT 9
ERBB_ALV
ID ERBB_ALV STANDARD; PRT; 634 AA.
AC P00534;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein erbb (EC 2.7.1.112).
GN V-ERBB.
OS Avian leukosis virus.
OC Viruses; Retroviruses; Retroviridae; Alpharetrovirus.
OC NCBI_TaxID=11864;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=85228222; PubMed=2988784;
RA Nilsen T.W., Maroney P.A., Goodwin R.G., Rottman F.M.,
RA Crittenden L.B., Raines M.A., Kung H.-U.;
RT "c-erbB activation in AlV-induced erythroblastosis: novel RNA
RT processing and promoter insertion result in expression of an

```

RT amino-truncated EGF receptor.
 RL Cell 41:719-726(1985).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS PRODUCED BY ERYTHROLEUKEMIA CELLS
 CC IN CHICKS WHICH HAVE BEEN INJECTED WITH THE AVIAN LEUKOSIS VIRUS
 CC AND SUBSEQUENTLY DEVELOP LEUKEMIA. THE LEUKEMIAS ARE INDUCED BY
 CC THE INSERTION OF RETROVIRAL PROMOTER ELEMENTS IN OR NEAR THE
 CC C-ERBB GENE, WHICH LEADS TO C-ERBB ACTIVATION.
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-ENV-ERBB
 CC PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: M10066; AAA48763.1; ALT_INIT.
 CC PIR: A00643; TVCHLV.
 CC PIR: B00643; TVFVLV.
 CC HSSP: P1352; 1FCR.
 CC InterPro: IPR00719; Euk_pkinase.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam: PF00069; pkinase; 1.
 CC PRINTS: PR00109; TYRKINASE.
 CC ProDom: PD000001; Euk_pkinase; 1.
 CC SMART: SM00219; TyrcK; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC KMW transferase: Tyrosine-protein kinase; ATP-binding; Oncogene;
 CC Glycoprotein; Phosphorylation.
 CC FT DOMAIN 132 399
 CC FT NP_BIND 138 146 ATP (BY SIMILARITY).
 CC FT BINDING 165 165 ATP (BY SIMILARITY).
 CC FT ACT_SITE 257 257 BY SIMILARITY.
 CC FT SEQUENCE 634 AA; 70891 MW; E705E33A0BE01FCC CRC64;
 SQ
 Query Match 11.7%; Score 138; DB 1; Length 634;
 Best Local Similarity 24.5%; Pred. No. 0.085;
 Matches 54; Conservative 25; Mismatches 59; Indels 82; Gaps 10;
 QY 35 RSPL-----APSEGAGSDVFDGLGMAKGLQSLPTHPDPSLPQRYSEDPVLPSPET-- 87
 DB 451 RFLLSLSLSTNSNSATNCID-----RNGGCHYVREDSFYQRTSSDPTGFLRESID 502
 QY 88 DGVAPLTCSPQPEYVNPDPVRFQPPSPREGPLPAARPAAGATLERAKTISPGKNGVAKDV 147
 DB 503 DGFLL-----PAPEYVNO--LMPKKPS-----TAMVNOQIYNNI 533
 QY 148 F-----AFGAVENPEYLTPOGGAAPQHPHPAFNFDLYWDQ----- 188
 DB 534 SLTAISKLPMDSKYQNSHSTAVDNPETL-----NTNOSPLAKTIVFESSPYWIOSGNH 565
 QY 189 ----DPPE-----RGAPPSTFKGTPTAENPEYGLDVP 217
 DB 586 QINLDNPDYQODFLPMETKPNGLKVPAAENPEYLAAP 625
 RESULT 10
 PRP1_HUMAN STANDARD: PRT: 331 AA.
 AC P04280:
 DT 20-MAR-1987 (rel. 04, Created)
 DT 13-AUG-1987 (rel. 05, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Salivary proline-rich protein precursor (Clones CP3, CP4 and CP5)
 DE [Contains: Basic peptide IB-6; Peptide P-H].
 GN PRB1.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85289325; PubMed=2993301;
 RA Maeda N., Kim H.-S., Azen E.A., Smithies O.;
 RT "Differential RNA splicing and post-translational cleavages in the
 RT human salivary proline-rich protein gene system.";
 RL J. Biol. Chem. 260:11123-11130(1985).
 RN [2]
 RP SEQUENCE OF 214-331.
 RX MEDLINE=86243355; PubMed=3521730;
 RA Kaufman D., Hofmann T., Bennick A., Keller P.;
 RT "Basic proline-rich proteins from human parotid saliva: complete
 RT covalent structures of proteins IB-1 and IB-6.";
 RL Biochemistry 25:2387-2392(1986).
 RN [3]
 RP SEQUENCE OF 276-331.
 RX MEDLINE=84161824; PubMed=6671974;
 RA Salton E., Isemura S., Sanada K.;
 RT "Further fractionation of basic proline-rich peptides from human
 RT parotid saliva and complete amino acid sequence of basic proline-rich
 RT peptide P-H.";
 RL J. Biochem. 94:1991-1997(1983).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: K03204; AAA60185.1; -
 CC DR EMBL: K03205; AAA60186.1; -
 CC DR EMBL: K03206; AAA60187.1; -
 CC DR PIR: A03291; PIH0B6.
 CC DR PIR: C25372; C25372.
 CC DR Genew: HGNC:9337; PRB1.
 CC MIM: 180989;
 CC KW Repeat; Parotid gland; Multigene family; Saliva; Signal.
 CC FT SIGNAL 1 16
 CC FT CHAIN 214 331 PEPTIDE IB-6.
 CC FT CHAIN 276 331 PEPTIDE P-H.
 CC FT VARIANT 106 238 MISSING (IN CLONE CP-4).
 CC FT VARIANT 106 238 /FTID-VAR.005561.
 CC FT VARIANT 106 258 MISSING (IN CLONE CP-5).
 CC FT CONFLICT 276 276 A -> S (IN REF. 2 AND 3).
 CC FT SEQUENCE 331 AA; 32596 MW; 3F481FE8BA39751 CRC64;
 SQ
 Query Match 11.5%; Score 135.5; DB 1; Length 331;
 Best Local Similarity 25.5%; Pred. No. 0.064;
 Matches 64; Conservative 17; Mismatches 76; Indels 94; Gaps 13;
 QY 11 RSSSTRSGGDLTLGLEPSEEAAPRSPLAP---SEG---AGSDVFDGLGMAKGLQSL 64
 DB 86 KSRSPRSPEPK-PQGRPPPGGNGPQGPGRPPPGKPGQSGKPPGPPKPGQSGK 140
 QY 65 PTHDPSPLQRYSEDPTVPPLPSETDGYVAPLTGSPQPEYVNPDPVRFQPPSPS 116
 DB 141 ----PPQGGKSGSPRSP-PGKPGQ-----PPQGGKQPGPGRPPPGKPGQSGK 187
 QY 117 ---EGPLPAARPAAGATLERAKTIS-----PGKNGVAVDVFAGAVENPEYLTPOGGAAP 169
 DB 188 NRPQGPPEPKPGQPPPGQSGKSPQSPPEK-----PGQPPQGGNQPG 231
 QY 170 PHPPPAFSPAFDLXYWDQDPERG-----APPSTFK 201
 DB 232 GPPPPPGKP-----QGPPQGGNKPQGPGRPPPGKPGQSGKSGSARAPGPKQ 282

QY 202 GTPAE--NPE 210
 Db 283 GPOQEGNNPQ 293

RESULT 11
 ODO2_MCTU STANDARD: PRT: 553 AA.

AC 010381:
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Dihydroliipoamide succinyltransferase component of 2-oxoglutarate
 dehydrogenase complex (EC 2.3.1.61) (E2).
 GN SUGB OR RV2215 OR MT2272 OR MTCY190.26.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV:
 RX MEDLINE=96295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock R., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Felkell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sultun J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleisichmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayan L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Knouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE 2-OXOGLUTARATE DEHYDROGENASE COMPLEX CATALYZES THE
 CC OVERALL CONVERSION OF 2-OXOGLUTARATE TO SUCCINYL-COA & CO(2). IT
 CC CONTAINS MULTIPLE COPIES OF 3 ENZYMATIC COMPONENTS: 2-OXOGLUTARATE
 CC DEHYDROGENASE (E1), DIHYDROLIPOAMIDE SUCCINYLTRANSFERASE (E2) AND
 CC LIPOAMIDE DEHYDROGENASE (E3) (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + dihydrolipoamide - CoA + S-
 CC succinylidihydrolipoamide.
 CC -1- COFACTOR: CONTAINS TWO COVALENTLY-BOUND LIPOYL COFACTOR
 CC (POTENTIAL).
 CC -1- PATHWAY: tricarboxylic acid cycle.
 CC -1- SUBUNIT: FORMS A 24-POLYPEPTIDE STRUCTURAL CORE WITH OCTAHEDRAL
 CC SYMMETRY (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE 2-OXOACID DEHYDROGENASE FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 LIPOYL-BINDING DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: Z70283; CA94256.1; -
 DR EMBL: AE007072; AAK6557.1; -
 DR HSSP: P07016; 1c4t.
 DR TIGR: MT2272; -
 DR TubercuList; RV2215; -

DR InterPro: IPR001078; 2oxoacid_dh.
 DR InterPro: IPR000089; biotin_lipoyl.
 DR InterPro: IPR004167; e3_binding.
 DR InterPro: IPR003016; lipoyl.
 DR Pfam: PF00198; 2-oxoacid_dh; 1.
 DR Pfam: PF00364; biotin_lipoyl; 2.
 DR Pfam: PF02817; e3_binding; 1.
 DR ProDom: PD001115; 2oxoacid_dh; 1.
 DR PROSITE: PS00189; lipoyl; 2.
 KW Tricarboxylic acid cycle; Transferase; Acyltransferase; Lipoyl;
 KW Complete proteome.
 FT BINDING 43 LIPOYL (POTENTIAL).
 FT BINDING 162 LIPOYL (POTENTIAL).
 FT ACT_SITE 523 BY SIMILARITY.
 FT ACT_SITE 527 BY SIMILARITY.
 SO SEQUENCE 553 AA; 57087 MW; 54B6E70D23B804A7 CRC64;

Query Match 11.5%; Score 135.5; DB 1; Length 553;
 Best Local Similarity 25.9%; Pred. No. 0.11;
 Matches 56; Conservative 24; Mismatches 75; Indels 61; Gaps 10;

QY 27 EPSEEARSP--APSEAGSDVFDGLGMAKG-----L 61
 Db 101 KPAEPPEPVQPTSGAPAGDAPKPYLMPELGESVTEGTVIRMLKKIGDSVGYDEPLVEYST 160
 QY 62 QSLPHDPSPLD-----RSEDPVPLPSEND--GYAVPLCSPOPEYVNOPTVRPOP- 112
 Db 161 DKVDTEIPVAGVLYSTISADBDATVPVGGELARIGVADIGAAPAPAPAP--VPEPA 218
 QY 113 PSPREGPL--PAAPAGA-----TLERAKTLPKNGVVK--DVFA 149
 Db 219 PTPKNEPAPSPAPAPAGAPVYVTLVKRLASENNIDLAGVGTGVGRIRKQDVLA 278
 QY 150 FCGAVENPEYLTP--QCGAAPQPPPPAPSPAFNL 183
 Db 279 AAQKRAKAPAPAAQAAAAAPAPAPAPAPALML 314

RESULT 12
 EXTN_TOBAC STANDARD: PRT: 620 AA.

AC P13983;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extensin precursor (Cell wall hydroxyproline-rich glycoprotein).
 GN HRCPTN3.
 OS Nicotiana tabacum (Common tobacco).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Xanthi; TISSUE=leaf;
 RX MEDLINE=90128263; PubMed=2612909;
 RA Keller B., Lamb C.J.;
 RT "Specific expression of a novel cell wall hydroxyproline-rich
 RT glycoprotein gene in lateral root initiation.";
 RL Genes Dev. 3:1639-1646(1989).
 CC -1- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN
 CC THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE
 CC MAIN ROOT.
 CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
 CC -1- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE
 CC SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
 CC GLYCOSYLATED.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

FT REPEAT 75 300 APPROXIMATE.
FT REPEAT 390 600 APPROXIMATE.
FT DOMAIN 1028 1071 SER-RICH.
FT NE_BIND 714 981 PROTEIN KINASE.
FT BINDING 720 728 ATP (BY SIMILARITY).
FT ACT_SITE 839 839 ATP (BY SIMILARITY).
FT DISULFID 190 199 BY SIMILARITY.
FT DISULFID 194 207 BY SIMILARITY.
FT DISULFID 215 223 BY SIMILARITY.
FT DISULFID 219 231 BY SIMILARITY.
FT DISULFID 232 240 BY SIMILARITY.
FT DISULFID 236 248 BY SIMILARITY.
FT DISULFID 251 260 BY SIMILARITY.
FT DISULFID 264 291 BY SIMILARITY.
FT DISULFID 295 307 BY SIMILARITY.
FT DISULFID 311 326 BY SIMILARITY.
FT DISULFID 329 333 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 526 535 BY SIMILARITY.
FT DISULFID 539 555 BY SIMILARITY.
FT DISULFID 558 571 BY SIMILARITY.
FT DISULFID 562 579 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT DISULFID 595 617 BY SIMILARITY.
FT DISULFID 620 628 BY SIMILARITY.
FT DISULFID 624 636 BY SIMILARITY.
FT MOD_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 528 528 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 19 19 C -> S (IN REF. 2).
FT CONFLICT 539 539 C -> W (IN REF. 5).
FT CONFLICT 991 991 L -> F (IN REF. 4).
FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;

```

Query Match 11.38; Score 134; DB 1; Length 1210;
 Best Local Similarity 26.48; Pred. No. 0.27; Mismatches 63; Indels 76; Gaps 11;
 Matches 58; Conservative 23;

```

QY 47 DVEDGDLGMAKGLQSLPTHPSP-----LQRYYS 76
  ||||| : : : : : |||||
DB 1011 DVDADAEYLIPQGFNFSPSTRPLLSLATSNNSTVACINNGSCRYKEDAFLOKYS 1070
  ||||| : : : : : |||||
QY 77 EDPIVPLPSEF--DGIYAPLTCSPQPEYVNPQVPRPSPREPLRAARAGATLERAK 134
  ||||| : : : : : |||||
DB 1071 SDIPGAVTEENIDAFI-----PVPEYVNO-SVPRKPAAGVQNPVYHNPQLHP----- 1117
  ||||| : : : : : |||||
QY 135 TLPSPKGVYKDVAFAGAVENPEYL--TPOGAARQPHRPAPFAPDNLKYWDO----- 188
  ||||| : : : : : |||||
DB 1118 --AGRLNUN--PHSNAGNPEYLNTAO-----PCLSSGNSPALMILOKGSHO 1164
  ||||| : : : : : |||||
QY 189 -----DP-----PARGAPSTFKGTPTAENPEYLGDPV 217
  ||||| : : : : : |||||
DB 1165 MSLDNPDYQODFEFFKTKPNCFKNG-PTAENAEYLRVAPR 1203

```

RESULT 14
 SM6B_MOUSE

```

ID SM6B_MOUSE STANDARD; PRT; 886 AA.
AC 054951:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 6b precursor (Semaphorin VIB) (Sema VIB) (Semaphorin N)
DE (Sema N).
GN SEMA6B OR SEMA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98027184; PubMed=9361278;
RA Eckhardt F., Behar O., Calautti E., Yonezawa K., Nishimoto I.,
RA Fishman M.C.;
RT "A novel transmembrane semaphorin can bind c-src.";
RL Mol. Cell. Neurosci. 9:409-419(1997).
CC -1- SUBUNIT: HOMODIMER. BINDS SPECIFICALLY THE SH3 DOMAIN OF THE
CC PROONCOGENE C-SRC.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: DURING DEVELOPMENT IT IS EXPRESSED IN
CC MUSCLE. IN ADULTHOOD, IT IS EXPRESSED UBICUOUSLY.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF036585; AAC00493.1;
CC MGD: MGI:1202889; Sema6b.
CC Interpro: IPR003659; Plexin-like.
CC Interpro: IPR001627; Sema.
CC DR SMART: PF01403; Sema; 1.
CC DR SMART: SM00423; PSI; 1.
CC KW Signal; Transmembrane; Multigene family; Neurogenesis; Glycoprotein;
CC development; protein.
CC KM SIGNAL 1 26
CC FT CHAIN 27 886 POTENTIAL.
CC FT DOMAIN 27 605 SEMAPHORIN 6B.
CC FT TRANSMEM 606 626 EXTRACELLULAR (POTENTIAL).
CC FT DOMAIN 627 886 POTENTIAL.
CC FT DOMAIN 239 549 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 751 754 SEMA.
CC FT CARBOHYD 75 75 POLY-LEU.
CC FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 886 AA; 95466 MW; E5F56D125CDA57AD CRC64;

```

Query Match 11.08; Score 130.5; DB 1; Length 886;
 Best Local Similarity 25.88; Pred. No. 0.33; Mismatches 93; Indels 59; Gaps 10;
 Matches 61; Conservative 23;

```

QY 1 GAGG-----VWHRHRSSTRSGGDLTGLTSESEAP-----RSPLAPS 41
  ||||| : : : : : |||||
DB 669 GAGGPGGPPALLAPLQNKQNTAAALHGGPHDLDTGLTPTQTPLPQGRPLTPHPHA 728
  ||||| : : : : : |||||
QY 42 EGAGSDVFDGDLGMAKGLQSLPTHPSPPLQRYSEDPTVPL-PSYDGIYAPLTC----- 96
  ||||| : : : : : |||||
DB 729 HALGSRAMDHSHALLSASASTLLLAFA---RASEGQYVAFQEPESRLCAPSCRAHS 785
  ||||| : : : : : |||||
QY 97 -----SPOPEYVNPQVPRPQPSPRPG-PLPAARPAAGATLERAKTLPQKNG--- 142
  ||||| : : : : : |||||

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:23 ; Search time 6.04097 Seconds
(without alignments)
1984.228 Million cell updates/sec

Title: US-09-821-883-23
Perfect score: 1587
Sequence: 1 STGYCTGDMKRLRPASPET.....FGASCVTACPYNYLSTDVGS 289

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1587	100.0	1255	1	ERRB2_HUMAN
2	1353.5	85.3	1257	1	ERRB2_RAT
3	1322	83.3	1254	1	ERRB2_MESAU
4	739	46.6	703	1	EGFR_CHICK
5	716.5	45.1	1210	1	EGFR_HUMAN
6	716.5	45.1	1210	1	EGFR_MOUSE
7	716.5	45.1	1308	1	ERRB4_HUMAN
8	706.5	44.5	1308	1	ERRB4_RAT
9	686	43.2	1342	1	ERRB3_HUMAN
10	658	41.5	1339	1	ERRB3_RAT
11	640.5	40.4	1167	1	XMRRK_XIPMA
12	535	33.7	1426	1	EGFR_DROME
13	396.5	25.0	1323	1	LRT3_CAEEL
14	315.5	19.9	1363	1	ILPR_BRALA
15	273.5	17.2	1477	1	HTK7_HYDAT
16	255	16.1	2146	1	INSR_DROME
17	246.5	15.5	1300	1	INSR_MOUSE
18	246.5	15.5	1607	1	IRR_CAVPO
19	245.5	15.5	1300	1	IRPR_LYMAST
20	244.5	15.4	1382	1	IRR_MOUSE
21	244.5	15.4	1383	1	INSR_HUMAN
22	243.5	15.3	1372	1	INSR_RAT
23	238.5	15.0	1297	1	INSR_MOUSE
24	232.5	14.7	1367	1	IRR_HUMAN
25	232.5	14.7	1370	1	IGIR_HUMAN
26	230.5	14.5	1373	1	IGIR_RAT
27	228.5	14.4	1390	1	IGIR_MOUSE
28	207	13.0	1390	1	INSR_AEDAE
29	140	8.8	1876	1	PKC5_BRACL
30	125	7.9	1875	1	PKC5_MOUSE
31	121	7.6	1680	1	PKC5_MOUSE
32	108.5	6.8	798	1	ITB0_DROME
33	106	6.7	1877	1	PKC5_RAT

34	105.5	6.6	798	1	ITB1_XENLA	P12606 xenopus lae
35	105	6.6	830	1	SREC_HUMAN	Q14162 homo sapien
36	104	6.6	937	1	PAC4_RAT	Q63415 rattus norv
37	103.5	6.5	769	1	ITB2_HUMAN	P05107 homo sapien
38	101.5	6.4	484	1	LEM2_PIG	P08110 sus scrofa
39	101.5	6.4	969	1	PAC4_HUMAN	P29122 homo sapien
40	100	6.3	551	1	LEM2_RABIT	P27113 oryctolagus
41	98	6.2	3075	1	LMAI_HUMAN	P25391 homo sapien
42	97	6.1	837	1	ATSA_HUMAN	C75171 homo sapien
43	97	6.1	3106	1	LMMA_MOUSE	Q63675 mus musculu
44	96.5	6.1	2282	1	ZAN_RABIT	P57999 oryctolagus
45	96	6.0	667	1	TS11_GIALA	Q03185 giardia lam

ALIGNMENTS

RESULT 1	ID	ERRB2_HUMAN	STANDARD:	PRT: 1255 AA.
AC	P04626:			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC:2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell			
DE	surface receptor HER2) (MIM 19).			
CN	ERRB2 OR HER2 OR NGL OR NEU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86118663; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,			
RA	Saito T., Toyoshima K.;			
RT	"Similarity of protein encoded by the human c-erbB-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86070181; PubMed=2999974;			
RA	Consensus L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger T.,			
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
RN	[3]			
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=2995967;			
RA	Semba K., Kamata N., Toyoshima K., Yamamoto T.;			
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a			
RT	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
RN	[4]			
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	-1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX.			
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A			
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-			
CC	ALPHA AND AMBIREGULIN.			
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein			
CC	tyrosine phosphate.			
CC	-1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS			
CC	(POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			

RX MEDLINE-91222560; PubMed-2025425;
 RA Lai C., Lemke G.;
 RT "an extended family of protein-tyrosine kinase genes differentially
 RT expressed in the vertebrate nervous system.";
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE-92155181; PubMed-1346763;
 RA Gullick W.J., Bottomley A.C., Lofte F.J., Doak D.G., Mulvey D.,
 RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-
 RT oncogenic and oncogenic forms of the neu protein.";
 RL EMBO J. 11:43-48(1992).
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX.
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE, GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND ANPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.
 CC THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PFM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X03362; CA27059.1; ALT_INIT.
 DR PIR: A24562; TVRTNU.
 DR HSP: P11362; IFKG.
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002174; Puriin-like.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; Pkinase.1.
 DR Pfam: PF00757; Puriin-like.1.
 DR Pfam: PF01030; Recep_L.domain.2.
 DR Pfam: PF02757; YLP.2.
 DR ProDom: PD000001; Euk_Pkinase.1.
 DR SMART: SM00261; FU.3.
 DR SMART: SM00219; TYRK3.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 655 677 POTENTIAL.
 FT DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 159 369 CYS-RICH.
 FT DOMAIN 473 646 CYS-RICH.
 FT DOMAIN 722 989 PROTEIN KINASE.
 FT NP_BIND 728 736 ATP (BY SIMILARITY).
 FT BINDING 755 755 ATP (BY SIMILARITY).
 FT ACT_SITE 847 847 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 200 213 BY SIMILARITY.
 FT DISULFID 221 228 BY SIMILARITY.
 FT DISULFID 225 236 BY SIMILARITY.
 FT DISULFID 237 245 BY SIMILARITY.
 FT DISULFID 241 253 BY SIMILARITY.
 FT DISULFID 256 265 BY SIMILARITY.

FT DISULFID 269 296 BY SIMILARITY.
 FT DISULFID 300 312 BY SIMILARITY.
 FT DISULFID 316 332 BY SIMILARITY.
 FT DISULFID 335 339 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 517 530 BY SIMILARITY.
 FT DISULFID 533 542 BY SIMILARITY.
 FT DISULFID 546 562 BY SIMILARITY.
 FT DISULFID 565 578 BY SIMILARITY.
 FT DISULFID 569 586 BY SIMILARITY.
 FT DISULFID 589 598 BY SIMILARITY.
 FT DISULFID 602 625 BY SIMILARITY.
 FT DISULFID 628 636 BY SIMILARITY.
 FT DISULFID 632 644 BY SIMILARITY.
 FT MOD_RES 1141 1141 BY SIMILARITY.
 FT MOD_RES 1250 1250 BY SIMILARITY.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64.
 Query Match 85.3%; Score 1353.5; DB 1; Length 1257;
 Best Local Similarity 85.5%; Pred. No. 57e-107;
 Matches 247; Conservative 13; Mismatches 28; Indels 1; Gaps 1;
 QY 2 TQVCTGDMKRLRLPASPEHIDMLRHLYGCGVQVGNLELYLPTNLSLFLADTQIEVOG 61
 DB 23 TQVCTGDMKRLRLPASPEHIDMLRHLYGCGVQVGNLELYLPTNLSLFLADTQIEVOG 82
 QY 62 YVLAHNVGRVPRQRLRIYRGCTQLEFEDNVLAVIDNDPLNNTPTV-GASPGGLREIQ 120
 DB 83 YMLAHNVGRVPRQRLRIYRGCTQLEFEDKVALAVIDNDPLNNTPTV-GASPGGLREIQ 142
 QY 121 LRLSTLEIKGVLVLIQNPOLCYOPTILMKDIFHKNQIATLTLIDTNSRACHPCSPCKG 180
 DB 143 LRLSTLEIKGVLVLIQNPOLCYODMLKDYFRKNNQIAPVDIDTNSRACHPCSPCKG 202
 QY 181 SRGWESSEDCQSLRTFYVACGACARCKGRLPTDCCHQCAAGCTGPKHSDCLAHFNHS 240
 DB 203 NHCWESSEDCQSLRTFYVACGACARCKGRLPTDCCHQCAAGCTGPKHSDCLAHFNHS 262
 QY 241 GICELHCPALVTYNDTFESHPNPEGRTFGASCVTACPYVLTSDVCS 289
 DB 263 GICELHCPALVTYNDTFESHPNPEGRTFGASCVTACPYVLTSDVCS 311
 RESULT 3
 ERBB2_MESAU STANDARD; PRT: 1254 AA.
 ID ERBB2_MESAU
 AC Q60553;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbb-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
 GN ERBB2 OR NEU.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Mesocricetus.
 OX NCBI_Taxid=10036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Nerve;
 RX MEDLINE-94193007; PubMed-7908275;
 RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
 RA Yamazaki Y., Ishikawa T.;
 RT "Cloning and activation of the Syrian hamster neu proto-oncogene.";
 RL Gene 140:251-255(1994).

CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
 CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
 CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
 CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D16295; BAA03801.1; -
 CC HSSP: P11362; 1FGK.
 DR InterPro: IPR000494; EGFR_Ldomain.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; Pkinase: 1.
 DR Pfam: PF00757; Furin-like: 1.
 DR Pfam: PF01030; Recep_Ldomain: 2.
 DR Pfam: PF02757; YLP: 2.
 DR ProDom: PD000001; Euk_Pkinase: 1.
 DR SMART: SM00261; Fu: 3.
 DR SMART: SM00219; TYRKC: 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR: 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM: 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation.
 FT SIGNL 1 21 POTENTIAL.
 FT CHAIN 22 1254 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 653 675 POTENTIAL.
 FT DOMAIN 158 1254 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 472 644 CYS-RICH.
 FT DOMAIN 720 987 CYS-RICH.
 FT NP_BIND 726 734 ATP (BY SIMILARITY).
 FT BINDING 753 753 ATP (BY SIMILARITY).
 FT ACT_SITE 845 845 BY SIMILARITY.
 FT DISULFID 195 204 BY SIMILARITY.
 FT DISULFID 199 212 BY SIMILARITY.
 FT DISULFID 236 244 BY SIMILARITY.
 FT DISULFID 240 252 BY SIMILARITY.
 FT DISULFID 255 264 BY SIMILARITY.
 FT DISULFID 268 295 BY SIMILARITY.
 FT DISULFID 299 311 BY SIMILARITY.
 FT DISULFID 315 331 BY SIMILARITY.
 FT DISULFID 334 338 BY SIMILARITY.
 FT DISULFID 511 520 BY SIMILARITY.
 FT DISULFID 515 528 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 544 560 BY SIMILARITY.
 FT DISULFID 563 576 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 600 623 BY SIMILARITY.
 FT DISULFID 626 634 BY SIMILARITY.
 FT DISULFID 630 642 BY SIMILARITY.
 FT MOD_RES 1139 1139 BY SIMILARITY.
 FT MOD_RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1247 1247 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 571 571 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 629 629 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 658 658 V -> E (IN ONCOGENIC NEU).
 FT VARIANT 659 659 V -> E (IN ONCOGENIC NEU).
 SQ SEQUENCE 1254 AA; 138252 MM; 974C3791C21E2BE1 CRC64:
 Query Match 83.3%; Score 1322; DB 1; Length 1254;
 Best Local Similarity 83.3%; Pred. No. 2,7e-104;
 Matches 240; Conservative 16; Mismatches 32; Indels 0; Gaps 0;
 Oy 2 TQVCTGDMKRLRPASPEHLDMLRHLYOCQVQVQVGLTELYPTNLSFLDIOEYOG 61
 Db 23 TQVCTGDMKRLRPASPEHLDMLRHLYOCQVQVQVGLTELYPTNLSFLDIOEYOG 82
 Oy 62 YVLIANOVROVPLQRLRIYRGTOLEFEDNYALVLDNGDPLNNTPYVGASPGGLRELOL 121
 Db 83 YVLIANOVROVPLQRLRIYRGTOLEFEDKYALAVLDNRDPLDNTVATGRTPEGLRELOL 142
 Oy 122 RSLTEILKGVLIQNRNQLCYOTIILMKDIFHKNNQALLLIDTNSRACHPSPCKSGS 181
 Db 143 RSLTEILKGVLIQNRNQLCYOTIILMKDIFHKNNQALPDIDTNSRACHPSPCKSGS 202
 Oy 182 RCMGESEDDQSILTRTYCAGCAGCARCKGLPTDCCHECCAGCTGPKISDCLAFHNSG 241
 Db 203 HCMGASPEDCOTLTGTITAPAVPAARARLPTDCCHECCAGCTGPKISDCLAFHNSG 262
 Oy 242 ICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTVGS 289
 Db 263 ICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTVGS 310
 RESULT 4
 ID EGF-CHICK STANDARD; PRT; 703 AA.
 AC P13387;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)
 DE (Fragment).
 GN EGF.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88261272; PubMed=3260329;
 RA Ullrich A., Vennstrom B., Schlessinger J., Givol D.;
 RT "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,
 RT expression in mouse cells, and differential binding of EGF and
 RT transforming growth factor alpha.";
 RT Mol. Cell. Biol. 8:1970-1978(1988).
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

FT DISULFID 558 571 BY SIMILARITY.
 FT DISULFID 562 579 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT DISULFID 595 617 BY SIMILARITY.
 FT DISULFID 620 628 BY SIMILARITY.
 FT DISULFID 624 636 BY SIMILARITY.
 FT MOD_RES 680 680 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 FT MOD_RES 1092 1092 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1110 1110 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1172 1172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1197 1197 PHOSPHORYLATION (AUTO-, MAJOR SITE) (BY SIMILARITY).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 528 528 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 19 19 C -> S (IN REF. 2).
 FT CONFLICT 539 539 C -> W (IN REF. 5).
 FT CONFLICT 991 991 L -> F (IN REF. 4).
 FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
 SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;

Query Match 45.1%; Score 716.5; DB 1; Length 1210;
 Best local Similarity 47.3%; Pred. No. 5.4e-53;
 Matches 138; Conservative 39; Mismatches 96; Indels 19; Gaps 4;

QY 3 QVCTGDMKRLRASPETHLDMRLHYCGGVQGNELTYLPNMSLSLFDIOIQVQY 62
 DB 29 KVOGGSNRLTQJGFEHFLSLQRMVNNCEVVLGNLEITYVORNDLSLKTIOEVAGY 88
 QY 63 VLIHHNOVROVFORLRIVRQLEFEDNALAVLNDGDPINNTPTVYGASPGGIREQLR 122
 DB 89 VLIHHNOVROVFORLRIVRQLEFEDNALAVLNDGDPINNTPTVYGASPGGIREQLR 122
 QY 123 SLTEILKGVLLIQRNPOLCYQDTILMKDI---FHKNQALTLIDTNRSRACHPSPMC 178
 DB 139 NLOEILIGAVRFNSNPILTCNMOTIDRDYQVFMNSMDL---QSHRSPCKOPSC 194
 QY 179 KGRCKESEDGOSLTRYCAGGCA-RCKGRLPTDCCHQCAAGCTGPRHSDCLACLF 237
 DB 195 PNGSGGGGEGNCKLRIKCAOCCSRGSRSPDCCHQCAAGCTGPRHSDCLACLF 237
 QY 238 NHSGICELHPALVTYNTDTFESMPNPEGRTGASCVTACPYNYLSTDVGS 289
 DB 255 QDNTKCTCPPLMLYNTTYQNDVNEGRYSFGATCVKCPKRYVYVDHGS 306

RESULT 7
 ID ERBA_HUMAN STANDARD; PRT: 1308 AA.
 AC Q15303;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)
 GN (p180erbB4) (Tyrosine kinase-type cell surface receptor HER4).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM JM-A).
 RC TISSUE-Breast carcinoma;
 RX MEDLINE=93189574; PubMed=838326;
 RA Plozman G.D., Culouscou J.-M., Whitney G.S., Green J.M., Carlton G.W.,
 Foy L., Neubauer M.G., Shoyab M.;

RT "Ligand-specific activation of HER4/p180erbB4, a fourth member of the
 RT epidermal growth factor receptor family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).
 RC TISSUE-Fetal brain;
 RX MEDLINE=97476287; PubMed=9334263;
 RA Elenius K., Corfas G., Paul S., Choi C.J., Rio C., Plozman G.D.,
 RA Klagsbrun M.;
 RT "A novel transmembrane domain isoform of HER4/erbB4. Isoform-specific
 RT tissue distribution and differential processing in response to
 RT phorbol ester.";
 RL J. Biol. Chem. 272:26761-26768(1997).
 CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
 CC 2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND
 CC NTAF. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.
 CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein.
 CC tyrosine phosphate.
 CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
 CC RECEPTORS (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER
 CC FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED
 CC BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND
 CC NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN
 CC CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,
 CC KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,
 CC PITUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,
 CC LUNG, SALIVARY GLAND, AND PANCREAS.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).

CC EMBL: L07868; AAB59446.1; -;
 DR HSSP: P11362; IFGK.
 DR Genew: HGNC:3432; ERBB4.
 DR MIM: 600543; -;
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-1ike.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; pkinase.1.
 DR Pfam: PF00757; Furin-1ike.1.
 DR Pfam: PF01030; Recep_L.domain.2.
 DR Pfam: PF02757; YLP.2.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR SMART: SM00261; FU.4.
 DR SMART: SM00219; TYRK.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Alternative splicing.
 FT SIGNAL 1 25
 FT CHAIN 26 1308
 FT DOMAIN 26 651
 FT TRANSMEM 652 675
 FT DOMAIN 676 1308
 FT DOMAIN 186 334
 FT DOMAIN 496 633
 FT CYT-PLAS 186 334
 FT CYT-RICH.
 FT CYS-RICH.
 FT DOMAIN 496 633
 FT CYS-RICH.

```

FT DOMAIN 718 985 PROTEIN KINASE.
FT NP_BIND 724 732 ATP (BY SIMILARITY).
FT BINDING 751 751 ATP (BY SIMILARITY).
FT ACT_SITE 843 843 BY SIMILARITY.
FT DISULFID 189 197 BY SIMILARITY.
FT DISULFID 193 205 BY SIMILARITY.
FT DISULFID 213 221 BY SIMILARITY.
FT DISULFID 217 229 BY SIMILARITY.
FT DISULFID 230 238 BY SIMILARITY.
FT DISULFID 234 246 BY SIMILARITY.
FT DISULFID 249 258 BY SIMILARITY.
FT DISULFID 262 289 BY SIMILARITY.
FT DISULFID 293 304 BY SIMILARITY.
FT DISULFID 308 323 BY SIMILARITY.
FT DISULFID 326 330 BY SIMILARITY.
FT DISULFID 503 512 BY SIMILARITY.
FT DISULFID 507 520 BY SIMILARITY.
FT DISULFID 523 532 BY SIMILARITY.
FT DISULFID 536 552 BY SIMILARITY.
FT DISULFID 555 569 BY SIMILARITY.
FT DISULFID 559 577 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 593 614 BY SIMILARITY.
FT DISULFID 617 625 BY SIMILARITY.
FT DISULFID 621 633 BY SIMILARITY.
FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).
FT VARSPLIC 626 648 NGPISHDCTTPWGHSTLPQHA -> IGSSIEDCIGLMD
      (IN ISOFORM JM-B).
SQ SEQUENCE 1308 AA: 146807 MW: 56480985D88761 CRC64:
Query Match 45.1%; Score 716.5; DB 1; Length 1308;
Best Local Similarity 48.6%; Pred. No. 5.9e-53;
Matches 138; Conservative 37; Mismatches 98; Indels 11; Gaps 3;

```

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
GN ERBB4 OR TYRO-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid-10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=98221155; PubMed=9553078;
RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
RA Marchionni M.A., Kelly R.A.;
RT "Neuregulins promote survival and growth of cardiac myocytes.
RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT ventricular myocytes."
RL J. Biol. Chem. 273:10261-10269(1998).
RN [2]
RP SEQUENCE OF 848-901 FROM N.A.
RC TISSUE=Sciatic nerve;
RX MEDLINE=91222560; PubMed=2025425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system."
RL Neuron 6:691-704(1991).
RN [3]
RP SEQUENCE OF 1031-1198 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spinal cord;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Froehner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuregulins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration."
RL J. Neurosci. 17:1642-1659(1997).
RN [4]
RP FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-
RN NTRK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION,
RN NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
CC RECEPTORS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING
CC NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS
CC OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE
CC RETICULAR NUCLEUS OF THE THALAMUS. VERY LOW LEVELS IN KIDNEY, AND
CC HEART.
CC -1- LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF041838; AAD08899.1;
DR EMBL: U52531; AAC53051.1;
DR HSP: P11362.1RGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_Pkinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.

```

```

DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00219; TYKIC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 25
FT CHAIN 26 1308
FT DOMAIN 26 651
FT TRANSMEM 652 675
FT DOMAIN 676 1308
FT DOMAIN 186 334
FT DOMAIN 496 633
FT DOMAIN 718 985
FT NP_BIND 724 732
FT BINDING 751 751
FT ACT_SITE 843 843
FT DISULFID 189 197
FT DISULFID 193 205
FT DISULFID 213 221
FT DISULFID 217 229
FT DISULFID 230 238
FT DISULFID 234 246
FT DISULFID 249 258
FT DISULFID 262 289
FT DISULFID 293 304
FT DISULFID 308 323
FT DISULFID 326 330
FT DISULFID 303 312
FT DISULFID 507 520
FT DISULFID 536 552
FT DISULFID 555 569
FT DISULFID 559 577
FT DISULFID 580 589
FT DISULFID 617 625
FT DISULFID 621 633
FT MOD_RES 1162 1162
FT MOD_RES 1188 1188
FT MOD_RES 1258 1258
FT MOD_RES 1284 1284
FT CARBOHYD 138 138
FT CARBOHYD 174 174
FT CARBOHYD 253 253
FT CARBOHYD 358 358
FT CARBOHYD 410 410
FT CARBOHYD 473 473
FT CARBOHYD 495 495
FT CARBOHYD 548 548
FT CARBOHYD 576 576
FT CARBOHYD 620 620
FT CONFLICT 1062 1062
FT CONFLICT 1080 1082
SQ SEQUENCE 1308 AA; 146957 MW; D944B0996A08B41 CRC64;

Query Match 44.5%; Score 706.5; DB 1; Length 1308;
Best Local Similarity 48.2%; Pred. No. 4, 2e-52;
Matches 137; Conservative 37; Mismatches 99; Indels 11; Gaps 3;

```

```

DB 136 LKMLTELLNGVYVDONKFLCYADTTHMDIVRNPMPSNMTLVSTIGSGCGRCHSCTG 195
OY 181 SRCWGESSEDCQSLTRTVACAGC-ARCKGGLPTDCHEQCAAGCTGPKHSDCLACLHFN 239
DB 196 -RCMGEPENHCQTLTRTVACAGC-ARCKGGLPTDCHEQCAAGCTGPKHSDCLACLHFN 254
OY 240 SGICELHCPALVYNTDFESMPNPEGRYFGASCYACPYNL 283
DB 255 SGACVYQCQPTFVYNPFTFQLEHNFNKAYYGAFCVKKCPHNFV 298.

RESULT 9
ERB3_HUMAN
ID ERB3_HUMAN STANDARD; PRT; 1342 AA.
AC P21860;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)
DE (c-erbB3) (Tyrosine kinase-type cell surface receptor HER3).
GN ERB3 OR HER3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90083234; PubMed-2687875;
RA Kraus M.H., Issing W., Miki T., Popescu N.C., Aaronson S.A.;
RT "Isolation and characterization of ERB3, a third member of the
RT ERB/epidermal growth factor receptor family: evidence for
RT overexpression in a subset of human mammary tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9193-9197(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-90311312; PubMed-2164210;
RA Plozman G.D., Whitney G.S., Neubauer M.G., Green J.M., McDonald V.L.,
RA Todaro G.J., Shoyab M.;
RT "Molecular cloning and expression of an additional epidermal growth
RT factor receptor-related gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:4905-4909(1990).
RN [3]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE-Placenta;
RX MEDLINE-93282822; PubMed-7685162;
RA Katoh M., Yazaki Y., Sugimura T., Terada M.;
RT "c-erbB3 gene encodes secreted as well as transmembrane receptor
RT tyrosine kinase.";
RL Biochem. Biophys. Res. Commun. 192:1189-1197(1993).
CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEURIGULINS AND NTK.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (LONG FORM) AND
CC SECRETED (SHORT FORM).
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS, A LONG TRANSMEMBRANE FORM AND A
CC SHORT SECRETED FORM, EXIST DUE TO ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EPITHELIAL TISSUES AND BRAIN.
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE (BY SIMILARITY).
CC -1- DISEASE: OVEREXPRESSED IN A SUBSET OF HUMAN MAMMARY TUMORS.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sdb.ch/announce/)

```

CC or send an email to license@sls-sib.ch.
 CC -----
 DR EMBL: M29366; AAA35790.1; -
 DR EMBL: M34309; AAA35979.1; -
 DR EMBL: S61953; AAB26935.1; -
 DR PIR: A36223; A36223.
 DR HSSP: P11362; 1FCG.
 DR Genew: HGNC:3431; ERBB3.
 DR MIM: 190151; -
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR001274; Furin-like.
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR Pfam: PF00069; Pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR ProDom: PD000001; Euk_Pkinase; 1.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00219; TYRC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Alternative splicing.
 FT SIGNAL 1 19
 FT CHAIN 20 1342
 FT TRANS 20 643
 FT TRANS 644 664
 FT DOMAIN 665 1342
 FT DOMAIN 709 966
 FT NP_BIND 715 723
 FT BINDING 742 742
 FT ACT_SITE 834 834
 FT DISULFID 186 194
 FT DISULFID 190 202
 FT DISULFID 210 218
 FT DISULFID 214 226
 FT DISULFID 227 235
 FT DISULFID 231 243
 FT DISULFID 246 255
 FT DISULFID 259 286
 FT DISULFID 290 301
 FT DISULFID 305 320
 FT DISULFID 323 327
 FT DISULFID 500 509
 FT DISULFID 504 517
 FT DISULFID 520 529
 FT DISULFID 533 549
 FT DISULFID 552 565
 FT DISULFID 556 573
 FT DISULFID 576 585
 FT DISULFID 589 610
 FT DISULFID 613 621
 FT DISULFID 617 629
 FT CARBOHYD 126 126
 FT CARBOHYD 250 250
 FT CARBOHYD 353 353
 FT CARBOHYD 408 408
 FT CARBOHYD 414 414
 FT CARBOHYD 437 437
 FT CARBOHYD 469 469
 FT CARBOHYD 522 522
 FT CARBOHYD 566 566
 FT CARBOHYD 616 616
 FT VANSPLIC 141 183
 FT VANSPLIC 184 1342
 FT VARSPLIC 560 560
 FT CONFLICT 1064 1064
 FT CONFLICT 1342 AA; 148097 MW; 7201EF66CA374BD CRC64;
 FT SEQUENCE

Query Match 43.2%; Score 686; DB 1; Length 1342;
 Best Local Similarity 46.1%; Pred. No. 2,46-50;
 Matches 131; Conservative 44; Mismatches 95; Indels 14; Gaps 5;
 QY 1 STQVCTGDMKRLPASEFTLMDIRLYGCGQVYQGLLELYLPTNLSLFDIDQEQ 60
 Db 25 SQAVCPGTLCNGSLVYGDENQYQTLTKLYERCEVVMGLIEYLGNMADLSLQWIRVET 84
 QY 61 GVLIVANQVQVPLQRIATVGTQLEFEDNYALVLDNGDPLNNTPTGASPGQLRELO 120
 Db 85 GVLIVANNEFTPLPLNLRVVGTOYDGKFAIFVM-----LNTVT-----NSSNALRQUR 135
 QY 121 LRSLEILKGGVLIQIRNPOLCYODTILMKDIFKKNQALFLIDTNRSPACSPCKG 180
 Db 136 LQGLLEILSGGYIEKNKLCIHMIDIRDLVROD---AEIVYDNGRSCPCHEVCKG 192
 QY 181 SRCWGESSEDCQSLRTVYACAGC-ARCKGRLPTDCHEQCAAGCTGPKHSQCLACIAPHNH 239
 Db 193 -RCWGPGESEDCQTLTKITCAFCQCNHCEGPNPNCCHDECAGCGSPDQDFACRHRND 251
 QY 240 SGICELHCPALVTYNTDFEESMPNPEGRVYTGASCVTACPNYL 283
 Db 252 SGACVPRCPPLVYKTLFQLEPNPHYTYQYGVGVASCAPHNFV 295

RESULT 10
 ERB3_RAT
 ID ERB3_RAT STANDARD: PRT: 1339 AA.
 AC Q62799; Q62955;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)
 DE (c-erbB3).
 GN ERB3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=96096535; PubMed=8522190;
 RA Hellyer N.J., Kim H.-H., Greaves C.H., Sierke S.L., Koland J.G.;
 RT "Cloning of the rat ErbB3 cDNA and characterization of the
 RT recombinant protein.";
 RL Gene 165:279-284(1995).
 RN [2]
 RP REVISIONS TO 85; 513 AND 565.
 RA Hellyer N.J., Koland J.G.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 922-1097 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Sciatic nerve;
 RX MEDLINE=97184212; PubMed=9030624;
 RA Carroll S.L., Miller M.L., Froehner P.W., Kim S.S., Corbett J.A.;
 RT "Expression of neurogins and their putative receptors, ErbB2 and
 RT ErbB3, is induced during Wallerian degeneration.";
 RL J. Neurosci. 17:1642-1659(1997).
 CC -1 FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAN.
 CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC -1 SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
 CC (POTENTIAL).
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1 DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
 CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
 CC -1 PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
 CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
 CC PHOSPHATIDYLINOSITOL-3-KINASE.
 CC -1 SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL: 029339; AAC28438.2; -
 DR EMBL: 052530; AAC3050.1; -
 DR HSSP: P11362; 1FGK.
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L.domain; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00261; FU; 5.
 DR SMART: SM00219; TYRKC; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
 KM Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 FT SIGNAL 1 19
 FT CHAIN 20 1339
 FT DOMAIN 20 643
 FT TRANSMEM 644 662
 FT DOMAIN 663 1339
 FT DOMAIN 183 259
 FT DOMAIN 707 964
 FT NP_BIND 713 721
 FT BINDING 740 740
 FT ACT_SITE 832 832
 FT DISULFID 186 194
 FT DISULFID 190 202
 FT DISULFID 210 218
 FT DISULFID 214 226
 FT DISULFID 227 235
 FT DISULFID 231 243
 FT DISULFID 246 255
 FT DISULFID 259 266
 FT DISULFID 290 301
 FT DISULFID 305 320
 FT DISULFID 323 327
 FT DISULFID 500 509
 FT DISULFID 504 517
 FT DISULFID 520 529
 FT DISULFID 533 549
 FT DISULFID 556 573
 FT DISULFID 576 585
 FT DISULFID 589 610
 FT DISULFID 613 621
 FT DISULFID 617 629
 FT DISULFID 126 136
 FT CARBOHYD 250 250
 FT CARBOHYD 353 353
 FT CARBOHYD 408 408
 FT CARBOHYD 414 414
 FT CARBOHYD 437 437
 FT CARBOHYD 459 459
 FT CARBOHYD 522 522
 FT CARBOHYD 566 566
 FT CARBOHYD 616 616
 FT CONFLICT 1028
 FT SEQUENCE 1339 AA; 147545 MW; 0AA5F2402BBDFLE CRC64;

Query Match 41.5%; Score 658; DB 1; Length 1339;
 Best Local Similarity 44.7%; Pred. No. 5; 6e-48;
 Matches 127; Conservative 42; Mismatches 101; Indels 14; Gaps 5;

QY 1 STGYCTGTDMKLRIPSPSEPHLDMLRHLVYGGCOVYGNLEFLTFLPTNASISFLDIOEVO 60
 DB 25 SQAVCEPTLNGLSGTGDADNOYOTFLKYRECEVMEIIVLTGHNADLSFLDMIREVT 84
 QY 61 GYVLIAHNOVROYPLQRLIRVGTQLFEDVYALAVLNDGNPLNNTPTVYASPGGJRELO 120
 DB 85 GYVLVAMNEFSVLPPLRVAVRGTOYVDGKFAIFVM-----LINTNT---NSSHALQOLK 135
 QY 121 LRSLETLKGVLIQRPOLCYODTILMKDIFKNNOLATLIDTNSRACHPCSPMKCG 180
 DB 136 FQLDTLTLSSGVYIEKNDKLCMDTIDMRIVNR---GAEIYVKNNGANCPCHEVCCKG 192
 QY 181 SRCWGESSEDCQSLTRTVACAGC-ARCKGPLPDDCHEGCAAGCTGPHSDCLACLFHFN 239
 DB 193 -RCWGPDPDCQLITKTIKICAPQCNCRGFGPNPQCCHEDCAGCSCGPDPTDCAFRFRND 251
 QY 240 SGICELHCPALVYVNDTFESMNPBERYTFGASCYVACRYNLT 283
 DB 252 SGACVPRCPPELVYVNLTLQLEPMPHTKYOGGVACVACPHNFV 295

RESULT 11
 ID XMRK_XIPMA STANDARD; PRT; 1167 AA.
 AC P13388;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Melanoma receptor protein-tyrosine kinase precursor (EC 2.7.1.112).
 GN XMRK OR TV.
 OS Xiphophorus maculatus (Southern platyfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
 OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
 OX NCBI_Taxid=8083;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90015140; PubMed-2797166;
 RA Witbrodt J., Adam D., Malitschek B., Mauelet W., Raulf F.,
 RA Telling A., Robertson S.M., Schartl M.;
 RT "Novel putative receptor tyrosine kinase encoded by the melanoma-
 RT inducing Tu locus in Xiphophorus.";
 RL Nature 341:415-421(1989).
 RN [2]
 RP REVISION TO 515.
 RA Schartl M.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLE RECEPTOR WITH PROTEIN-TYROSINE KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- DISEASE: INVOLVED IN PIGMENT CELLS MALIGNANT MELANOMAS.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR EMBL: X16891; CAA34770.2; -
 DR PIR: S06142; S06142.
 DR HSSP: P11362; 1FGK.
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.

```

DR Pfam: PF01030; Recep_L_domain: 2.
DR PRINTS: PRO0109; TYRKINASE.
DR ProDom: PD000001; Euk_Pkinase: 1.
DR SMART: SM00261; FU: 5.
DR SMART: SM00220; S_TKC: 1.
DR SMART: SM00219; TYRK: 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR: 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM: 1.
KM Transmembrane; Glycoprotein; Receptor; Signal; Transferase.
KT Tyrosine-protein kinase; ATP-binding; Phosphorylation; Proto-oncogene.
FT SIGNAL 1 25
FT CHAIN 26 1167
FT DOMAIN 26 642
FT TRANSMEM 643 665
FT DOMAIN 666 1167
FT DOMAIN 710 977
FT NE_BIND 716 724
FT BINDING 743 743
FT ACT_SITE 835 835
FT DISULFID 199 204
FT DISULFID 220 228
FT DISULFID 224 236
FT DISULFID 237 245
FT DISULFID 241 253
FT DISULFID 256 265
FT DISULFID 269 296
FT DISULFID 300 311
FT DISULFID 315 330
FT DISULFID 333 337
FT DISULFID 504 513
FT DISULFID 508 521
FT DISULFID 537 553
FT DISULFID 556 569
FT DISULFID 560 577
FT DISULFID 593 615
FT DISULFID 618 626
FT DISULFID 622 634
FT CARBOHYD 622 634
FT CARBOHYD 114 114
FT CARBOHYD 144 144
FT CARBOHYD 201 201
FT CARBOHYD 356 356
FT CARBOHYD 356 356
FT CARBOHYD 398 398
FT CARBOHYD 417 417
FT CARBOHYD 501 501
FT CARBOHYD 576 576
FT CARBOHYD 621 621
SQ SEQUENCE 1167 AA; 129934 MW; 4793E0749DC1D55A CRC64;

Query Match 40.4%; Score 640.5; DB 1; Length 1167;
Best Local Similarity 44.2%; Pred. No. 1.4e-66;
Matches 126; Conservative 40; Mismatches 112; Indels 7; Gaps 4;

```

```

QY 242 ICHELCPALVTYNTDFESHPNPEGRYTFGASCVTACPYNYLSTD 286
DB 264 TCKDTCPPEKTIYDIVSHQVVDNPNIKYTFGACVKECPSNVYVTE 308

RESULT 12
ID EGRF_DROME STANDARD: PRT; 1426 AA.
AC P04412; O61601; Q9W2G0; P81868;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Egfr)
DE (Gurken receptor) (Torpedo protein) (Drosophila relative of ERBB).
GN EGRF OR TOP OR C-ERBB OR DER OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
RX MEDLINE=94350209; PubMed=8070664;
RA Clifford R., Schupbach T.;
RT "Molecular analysis of the Drosophila EGF receptor homolog reveals
RT that several genetically defined classes of alleles cluster in
RT subdomains of the receptor protein.";
RL Genetics 137:531-550(1994).
RN [2]
RP REVISIONS.
RA Clifford R., Schupbach T.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=85124611; PubMed=2982499;
RA Livneh E., Glazer L., Segal D., Schlesinger J., Shilo B.-Z.;
RT "The Drosophila EGF receptor gene homolog: conservation of both
RT hormone binding and kinase domains.";
RL Cell 40:599-607(1985).
RN [4]
RP SEQUENCE FROM N.A. CHARACTERIZATION, AND ALTERNATIVE SPLICING.
RC STRAIN-Oregon-R; TISSUE=Embryo;
RX MEDLINE=87002474; PubMed=3093080;
RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
RT "Alternative 5' exons and tissue-specific expression of the
RT Drosophila EGF receptor homolog transcripts.";
RL Cell 46:1091-1101(1986).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM TYPE I), TISSUE SPECIFICITY, AND MUTATION
RP ANALYSIS.
RX MEDLINE=99102120; PubMed=9882502;
RA Lesokhin A.M., Yu S.-Y., Katz J., Baker N.E.;
RT "Several levels of EGF receptor signaling during photoreceptor
RT specification in wild-type, Ellipse, and null mutant Drosophila.";
RL Dev. Biol. 205:129-144(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadeu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

```


RA RA Doulbin K.J., Evangelista C.C., Ferraz N.S., Ferreira S., Fleischmann W.
RA Fodor C.G., Gabriellian A.E., Gang N.S., Gelbart P.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kechum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laiko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mlshina N.V., Moharry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclat J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert R., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye Y., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhang H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C. ;
RA "The genome sequence of *Drosophila melanogaster*." ;
RL Science 287:2185-2195(2000).
RL [7]
RP SEQUENCE OF 959-1078 FROM N.A.
RP STRAIN-Deekwanryeong ;
RC MEDLINE-95137938; Pubmed-2983232;
RX Madsenworth S.C., Vincent W.S. Iii, Bliedean-Wentworth D. ;
RA "A *Drosophila* genomic sequence with homology to human epidermal
RT growth factor receptor." ;
RL Nature 314:178-180(1985).
RL [8]
RN SEQUENCE OF 1133-1137, 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION
RP ANALYSIS.
RP MEDLINE-92038942; Pubmed-1936959 ;
RX Raz E., Schejter E.D., Shilo B. Z. ;
RA "interallelic complementation among DER/Flb alleles: implications for
RT the mechanism of signal transduction by receptor-tyrosine kinases." ;
RL Genetics 123:191-201(1991).
RN [9]
RP REVIEW.
RP MEDLINE-97248481; Pubmed-9094709 ;
RX Perrimon N., Perkins L.A. ;
RA "There must be 50 ways to rule the signal: the case of the *Drosophila*
RT EGF receptor." ;
RL Cell 89:13-16(1997).
CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,
CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAF-
CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.
CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES. AND FOR THE
CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL
CC POLARITIES OF THE OCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE
CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSEROSA
CC AND VENTRAL NEUROECTODERMAL CELLS. GERM BAND RETRACTION, CELL FATE
CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF
CC CUTICLE.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I
CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE
CC PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS, TYPE I (SHOWN HERE), TYPE II AND
CC TYPE III; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: UNBIOGENOUSLY EXPRESSED IN EMBRYOS. IN LARVAE,
CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENTAL DISK, ANLAGEN OF
CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST
CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH
CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN
CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX
CC AND THORACIC AND ABDOMINAL GANGLIA.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----

FT DISULFID 791 805 POTENTIAL.
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 551 551 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 746 746 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 776 776 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 368 368 C -> Y (IN SY10).
 FT VARIANT 469 469 C -> R (IN MN216).
 FT VARIANT 700 700 C -> W (IN MN23).
 FT VARIANT 753 753 C -> Y (IN SY11).
 FT VARIANT 1065 1065 G -> E (IN SY16).
 FT VARIANT 1074 1074 T -> I (IN SY16).
 FT VARIANT 1074 1074 G -> E (IN SY16).
 SQ SEQUENCE 1323 AA; 150510 MW; 680307EE53EEFA99 CRC64;

Query Match 25.0%; Score 396.5; DB 1; Length 1323;
 Best Local Similarity 28.8%; Pred. No. 8e-26;
 Matches 89; Conservative 54; Mismatches 113; Indels 53; Gaps 9;

OY 4 VCSTGDMKRLPASPETHLMDLRLHLYOGCGVGNLELYLPN----- 47
 DB 39 LCGSTNGISRYGTGNI-LEDELTWYRGCRVYGNLEITWIEANETKKMRESTNSYDPK 97
 OY 48 -----ASLSFLDIOEVGCVLIANOVROVLORLRYRGTQLEEDNALAVLNDGDP 101
 DB 98 NEDSPKSTINFNFNDLEIRSLIYRANIOKISFPRIRVYGEVPHDN-ALYIHNKDK- 155
 OY 102 LNTTPTVTGASPGCLRELOLRSLTEILKGVLLIQRNPOLCY-ODTLIMNDIFKNNQAL 160
 DB 156 -----VHEVYMRRLRVIRNGSVTTIQDNPKKCYIDKIDMKELTYDPP--VQ 199
 OY 161 TLIDTNRSRACH-----PCSPMKGSRGWSSESDCSLRTVYAGGCARC--KGPL 210
 DB 200 KYETTSNHOHCYONGKSMACHESC-NDKCSGSGDNDQGVYRSVCPKSCQCFYSNSTS 258
 OY 211 PTDCCHGCAAGCTGPKRHSCLACLFHNSGICELHCPALVTNTDTFESMPNREGRTYR 270
 DB 259 SYECCDSACGGCTGHPKNCIACSKYELDGIETCPSKRIRHGTGLRVFNPDRGYON 318
 OY 271 GASCVTACP 279
 DB 319 GNHCVEKCP 327

RESULT 14
 ILPR_BRALA
 ID ILPR_BRALA STANDARD: PRT: 1363 AA.
 AC 002466;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Insulin-like peptide receptor precursor (EC 2.7.1.112) (ILP receptor).
 OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.
 OC NCBI_Taxid=7740;
 OX [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=66408719; PubMed=8813726;
 RA Pashmiforush M., Chan S.J., Steiner D.F.:
 "Structure and expression of the insulin-like peptide receptor from
 RT amphioxus".
 RT Mol. Endocrinol. 10:857-866(1996).
 CC -1- FUNCTION: THIS RECEPTOR BINDS TO THE INSULIN RELATED PEPTIDE AND
 CC -1- HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBUNIT: PROBABLE Tetramer of 2 alpha and 2 beta chains linked by
 CC DISULFIDE BONDS. THE ALPHA CHAINS CONTRIBUTE TO THE FORMATION OF
 CC THE LIGAND-BINDING DOMAIN, WHILE THE BETA CHAIN CARRY THE KINASE
 CC DOMAIN (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN
 CC RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: S83394; AAB50848.1; -;
 DR HSSP: P06213; IIRK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003962; FNIII_repeat.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002011; RTK_naseII.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR PRINTS: PR00014; FNTYPEII.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00060; FN3; 3.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; ATP-binding; Phosphorylation; Signal.
 FT CHAIN 1 29
 FT SIGNAL 30 716
 FT PROPEP 717 720
 FT CHAIN 721 1363
 FT DOMAIN 721 928
 FT TRANSMEM 929 949
 FT DOMAIN 950 1363
 FT DOMAIN 994 1283
 FT NP_BIND 1000 1008
 FT BINDING 1028 1028
 FT ACT_SITE 1148 1148
 FT MOD_RES 1174 1174
 FT CARBOHYD 51 51
 FT CARBOHYD 97 97
 FT CARBOHYD 137 137
 FT CARBOHYD 278 278
 FT CARBOHYD 483 483
 FT CARBOHYD 599 599
 FT CARBOHYD 617 617
 FT CARBOHYD 665 665
 FT CARBOHYD 666 666
 FT CARBOHYD 711 711
 FT CARBOHYD 732 732
 FT CARBOHYD 736 736
 FT CARBOHYD 743 743
 FT CARBOHYD 816 816
 FT CARBOHYD 865 865
 FT CARBOHYD 898 898
 SQ SEQUENCE 1363 AA; 154104 MW; 238120BAEAB1ED55 CRC64;

Query Match 19.9%; Score 315.5; DB 1; Length 1363;
 Best Local Similarity 28.8%; Pred. No. 6.1e-19;
 Matches 92; Conservative 39; Mismatches 109; Indels 79; Gaps 16;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:49 ; Search time 19.3311 seconds
(without alignments)
3080.404 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGVCTGTDMLRLPASPETH.....FGASCYACPNYLTSDVGS 289

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: SP_archaea:*
3: SP_bacteria:*
4: SP_fungi:*
5: SP_human:*
6: SP_invertebrate:*
7: SP_mammal:*
8: SP_mhc:*
9: SP_organella:*
10: SP_phage:*
11: SP_plant:*
12: SP_rodent:*
13: SP_virus:*
14: SP_vertebrate:*
15: SP_unclassified:*
16: SP_virus:*
17: SP_bacteriophage:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1587	100.0	419 4 O9UK79	O9UK79 homo sapien
2	1450	91.4	1259 6 O18735	O18735 canis famill
3	739	46.6	527 13 O90836	O90836 gallus gall
4	723	45.6	149 6 O9BG66	O9BG66 oryctolagus
5	716.5	45.1	643 11 O9ERV6	O9ERV6 mus musculu
6	716.5	45.1	655 11 O9WVF5	O9WVF5 mus musculu
7	716.5	45.1	1210 11 O9EP98	O9EP98 mus musculu
8	708.5	44.6	1209 11 O9QX70	O9QX70 rattus norv
9	700.5	44.1	478 11 O9ESEO	O9ESEO rattus norv
10	686	43.2	331 4 O9BUD7	O9BUD7 homo sapien
11	649.5	40.9	1165 13 O9YH40	O9YH40 xiphophorus
12	635.5	40.0	599 13 O9PSH2	O9PSH2 gallus gall
13	596.5	37.6	1328 13 P79754	P79754 fugu rubrip
14	543	34.2	1433 5 O9BIH9	O9BIH9 anopheles g
15	407	25.6	1137 13 O9W6F6	O9W6F6 gallus gall
16	398	25.1	150 6 O9BG64	O9BG64 oryctolagus

17	395	24.9	151 6 O9BG65	O9BG65 oryctolagus
18	377	23.8	1368 5 O23821	O23821 caenorhabdi
19	341	21.5	366 5 O26569	O26569 schistosoma
20	341	21.5	1717 5 O26566	O26566 schistosoma
21	331	20.9	334 5 O26567	O26567 schistosoma
22	331	20.9	342 5 O26568	O26568 schistosoma
23	317	20.0	1193 5 O9YIX8	O9YIX8 ephydaria f
24	289	18.2	1671 5 O9NUV5	O9NUV5 blomphalaria
25	287.5	18.1	1472 5 O9U5A8	O9U5A8 bombyx mori
26	273.5	17.2	89 11 O88459	O88459 mus musculu
27	271.5	17.1	1418 13 O93457	O93457 scophthalmu
28	271.5	17.1	1418 13 O8UW83	O8UW83 paralicthys
29	268	16.9	1749 5 O8TOW6	O8TOW6 echinococcu
30	259.5	16.4	1412 13 O8UW84	O8UW84 paralicthys
31	256.5	16.2	1362 13 O9PV24	O9PV24 xenopus lae
32	255	16.1	2144 5 O9VD94	O9VD94 drosophila
33	252.5	15.9	1369 13 O8UW86	O8UW86 paralicthys
34	250	15.8	1358 13 O8UW85	O8UW85 xenopus lae
35	244	15.4	1368 13 O8UW85	O8UW85 paralicthys
36	232.5	14.7	1371 11 O9QVW4	O9QVW4 rattus norv
37	228.5	14.4	410 11 O63720	O63720 rattus norv
38	228.5	14.4	469 11 O63721	O63721 rattus norv
39	218.5	13.8	1245 13 O9VGH8	O9VGH8 scophthalmu
40	206.5	13.0	946 5 O9VJ04	O9VJ04 drosophila
41	205	12.9	868 5 O9VFE2	O9VFE2 drosophila
42	191	12.0	82 6 O9NOK4	O9NOK4 sus scrofa
43	189	11.9	131 5 O9BH16	O9BH16 anopheles g
44	185	11.7	1843 5 O968Y9	O968Y9 caenorhabdi
45	185	11.7	1846 5 O16131	O16131 caenorhabdi

ALIGNMENTS

RESULT 1

ID O9UK79 PRELIMINARY: PRT: 419 AA.
AC O9UK79;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Herstatin.
GN HER-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9415951; PubMed=10485918;
RA Doherty J.K., Bond C., Jardim A., Adelman J.P., Clinton G.M.;
RT "The HER-2/neu receptor tyrosine kinase gene encodes a secreted
autocatalytic inhibitor";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10869-10874(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Doherty J.K., Clinton G.M., Adelman J.P., Evans A.J., Hemner W.D.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF177761, AAD56009.2; -
DR InterPro: IPR000494; EGFRL-domain.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L-domain; 1.
DR SMART: SM00261; FU; 1
SO SMART: 419 AA; 45472 MW; FFC1BE347E2D030C CMC64;

Query Match 100.0%; Score 1587; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 2,6e-142;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 STGVCTGTDMLRLPASPETHLMDRLHYGCGVQVQGLLELYLPTNASISFLDIOEYQ 60
DB 22 STGVCTGTDMLRLPASPETHLMDRLHYGCGVQVQGLLELYLPTNASISFLDIOEYQ 81

Oy	61	GVLIAHQVQVOPVQJRIYRGQJLEDNADYALVLDNGPLNNTPTVTGASPEGLELO	120
Dd	82	GVLIAHQVQVOPVQJRIYRGQJLEDNADYALVLDNGPLNNTPTVTGASPEGLELO	141
Oy	121	LRSLTEILKGVLLIORNPQLCYODTILMKDIFHKNNQDLATLIDTNSRACHPCSPCKG	180
Dd	142	LRSLTEILKGVLLIORNPQLCYODTILMKDIFHKNNQDLATLIDTNSRACHPCSPCKG	201
Oy	181	SRMCWESSEDDOQLTRYCAGGCAKCKGRLPTDCCHEOCAAGCTGPRHSDCLACLFHNHS	240
Dd	202	SRMCWESSEDDOQLTRYCAGGCAKCKGRLPTDCCHEOCAAGCTGPRHSDCLACLFHNHS	261
Oy	241	GICELHCALVYNTDPESPMPNEGRTFCAACVTAACPYNVISTDYGS	289
Dd	262	GICELHCALVYNTDPESPMPNEGRTFCAACVTAACPYNVISTDYGS	310

DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Receptor tyrosine kinase ErbB2 (Fragment).
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
 RA Tetens F., Fischer B.
 RT "Erbb genes and epidermal growth factor- (EGF-) like ligands in the
 RT peti-implantation rabbit uterus and blastocyst."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF331378; AAK14371.1;
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR SMART: SM00261; FU; 2.
 KM Kinase.
 FT NON_TER
 FT 1 149
 SQ SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;
 Query Match 45.68; Score 723; DB 6; Length 149;
 Best Local Similarity 81.28; Pred. No. 7.5e-61;
 Matches 121; Conservative 8; Mismatches 20; Indels 0; Gaps 0;
 QY 137 NPQCYDTILMKDIFHNQALTLIDTNRSRACHPCSPKSGESSEDCQSLTR 196
 DB 1 NPQCYDTILMKDIFHNQALTLIDTNRSRACHPCSPKSGESSEDCQSLTR 60
 QY 197 TVCAGGACRCKGRLPTDCHEOCAGCTGPRHSDCLAGHNSGICELHCPALVTYNTD 256
 DB 61 TICAGGACRCKGRLPTDCHEOCAGCTGPRHSDCLAGHNSGICELHCPALVTYNTD 120
 QY 257 TFESEMPNEGRYTFGASCVTACPYNYLST 285
 DB 121 TFESEMPNEGRYTFGASCVTACPYNYLST 149
 RESULT 5
 Q9ERV6
 ID Q9ERV6 PRELIMINARY; PRT; 643 AA.
 AC Q9ERV6;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Epidermal growth factor receptor isoform 2.
 GN EGFR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-C3H/101, 129/SVJ, AND 129/SVEV7AC;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Maibhe N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF275366; AAG28046.1;
 DR EMBL; AF275365; AAG28046.1; JOINED.
 DR MGD; MGI:95294; Egfr.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000494; EGFR_Ldomain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recept_Ldomain; 2.
 DR SMART: SM00261; FU; 4.

DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 KM Receptor.
 SQ SEQUENCE 643 AA; 71476 MW; DEF22002C084911B1 CRC64;
 Query Match 45.18; Score 716.5; DB 11; Length 643;
 Best Local Similarity 47.38; Pred. No. 1.9e-59;
 Matches 138; Conservative 39; Mismatches 96; Indels 19; Gaps 4;
 QY 3 QVCTGTDMKLLRPA SPENHLDMLRHLYGCGVGVGNELTYLPNNASLPFLODIOEVGY 62
 DB 29 KVCQTSRLRLQGLTFEHPFLSLQRMNCEVJLGNLETTYVORNYDSLTKTIOEVAGI 88
 QY 63 VLIANVOYPLQRLRIVRGTOLFEEDNYVALVDNGPLNNTTPYVTCASPGRLQLR 122
 DB 89 VLIALNVERIPLENLQIRGNALYENTYALATISN-----YGTNLTGLRELPMR 138
 QY 123 SLTELKGVLIQRNPOLCYODTILMKDI---FHKNQALTLIDTNRSRACHPCSPMC 178
 DB 139 NLOETLIGAVRFSNNPILCNMDTIQWRDIYONVFNMSNMDL---OSHPSRCRCPSC 194
 QY 179 KGSRCWGESSEDCQSLTRTVACGGA--RCKGRLPTDCHEOCAGCTGPRHSDCLAGH 237
 DB 195 PNGSCWGGEGEENCOXKRLTKICAQC SHRCGRSPSDCCNOCAGCTGPRHSDCLAGH 254
 QY 238 NMSGICELHCPALVTYNTDFESMPNEGRYTFGASCVTACPYNYLSTVGS 289
 DB 255 QDEATCKRCKTCLPLMLYNTTYQMVONPEGKYSFGATCYKCKPRMYVTVYDHGS 306
 RESULT 6
 Q9WVF5
 ID Q9WVF5 PRELIMINARY; PRT; 655 AA.
 AC Q9WVF5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Epidermal growth factor receptor (Epidermal growth factor receptor
 DE isoform 3).
 GN EGFR.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-C57BL/6J; TISSUE=LIVER.
 RA Reiter J.L., Threadgill D.W., Danielson A.J., Scheil C.,
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
 RA Maibhe N.J.;
 RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode
 RT Carboxy-Terminal Truncated Receptors."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN-C3H/101, 129/SVJ, AND 129/SVEV7AC;
 RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Maibhe N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN-C57BL/6J; TISSUE=LIVER.
 RA MEDLINE-21085660; PubMed-11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Katsukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schmitt L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carinici P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mondraets P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Sessa T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RT
 RL Nucleus 409:685-690(2001).
 DR EMBL: AF124513: AAD44149.1; -
 DR EMBL: AF275366: AAG28047.1; JOINED.
 DR EMBL: AF275364: AAG28047.1; JOINED.
 DR EMBL: AF275365: AAG28047.1; JOINED.
 DR EMBL: AK004883: BAB23681.1; -
 DR EMBL: AK004883: BAB23681.1; -
 DR EMBL: AK004911: BAB23682.1; -
 DR MGD: MGI:95294; Egfr.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like: 1.
 DR SMART: SM01030; Recep_L_domain: 2.
 DR SMART: SM00261; FvJ: 3.
 KW RECEPTOR.
 SO SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;

Query Match 45.1%; Score 716.5; DB 11; Length 655;
 Best Local Similarity 47.3%; Pred. No. 1.9e-59;
 Matches 138; Conservative 39; Mismatches 96; Indels 19; Gaps 4;
 QY 3 QVCTGDMKRLPASPETHLDMRLHYQGVQGVQGNLELYLPNMSLSFLQDIOEVQY 62
 DB 29 KVGQSTNRRLTQGTGFEDHFLSLQRMYNCEVVLGNLEITYVQRNYSLSLKTIOEVAGY 88
 QY 63 VLIANNOVROVPLQRLRIYVGTQLFEDNVALAVLNDGDPNNTPVYGASPGSLRELQRL 122
 DB 89 VLIALNVERIPLENLQIINGNALYENTVALALSN-----YGTNRGTGLRELPMR 138
 QY 123 SLTEILKGVLIORNPOLCYODTIILMKDI----FKKNQALALFLIDNRSRACHPCSPMC 178
 DB 139 NLQELILGAVRFSNNPLTLCNMDTIOWRDIVQNVFMSNMSMDL-----QSHPSSCPCKDPCSC 194
 QY 179 KGSRCWSESESDOSLTRYACAGCA-RCKGRLPTDCHRCQCAAGCTGPKHSCLACLHF 237
 DB 195 PNGSCWGGEGENCOKLKIICAGQCSHRGSRSPDCCHNQCAGCTGPRESDCLVQCKF 254
 QY 238 NMSGICELHCPALVYNTDTFESMPNPEGRYTGASCVTACPYNYLSTDVGS 289
 DB 255 QDEATCKDTCPLMLNPTTYQMDVNEPKYSGATCVKRCPRNYVTVDHGS 306

RESULT 7
 Q9EP98 PRELIMINARY; PRT: 1210 AA.
 AC Q9EP98;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Epidermal growth factor receptor isoform 1.
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVAC;
 RA Reller J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Pearsall R.S., Green P.J., Yee D., Lampand A.L.,

RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Wahl N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RA Reller J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Soehl C., Pearsall R.S., Green P.J., Yee D., Lampand A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Wahl N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egfr transcripts encoding truncated receptor
 RT isoforms.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275366: AAG28045.1; -
 DR EMBL: AF275364: AAG28045.1; JOINED.
 DR EMBL: AF275365: AAG28045.1; JOINED.
 DR EMBL: AF275367: AAG24386.1; -
 DR HSSP: P11362; 1FGK.
 DR MGD: MGI:95294; Egfr.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00757; Furin-like: 1.
 DR Pfam: PF00069; pkinase: 1.
 DR Pfam: PF01030; Recep_L_domain: 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_pkinase: 1.
 DR SMART: SM00220; S_TKc: 1.
 DR SMART: SM00261; FvJ: 5.
 DR SMART: SM00219; TyrcKc: 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Receptor; Transferase.
 SO SEQUENCE 1210 AA; 134840 MW; 62CDD021C9DB32E18 CRC64;

Query Match 45.1%; Score 716.5; DB 11; Length 1210;
 Best Local Similarity 47.3%; Pred. No. 4e-59;
 Matches 138; Conservative 39; Mismatches 96; Indels 19; Gaps 4;
 QY 3 QVCTGDMKRLPASPETHLDMRLHYQGVQGVQGNLELYLPNMSLSFLQDIOEVQY 62
 DB 29 KVGQSTNRRLTQGTGFEDHFLSLQRMYNCEVVLGNLEITYVQRNYSLSLKTIOEVAGY 88
 QY 63 VLIANNOVROVPLQRLRIYVGTQLFEDNVALAVLNDGDPNNTPVYGASPGSLRELQRL 122
 DB 89 VLIALNVERIPLENLQIINGNALYENTVALALSN-----YGTNRGTGLRELPMR 138
 QY 123 SLTEILKGVLIORNPOLCYODTIILMKDI----FKKNQALALFLIDNRSRACHPCSPMC 178
 DB 139 NLQELILGAVRFSNNPLTLCNMDTIOWRDIVQNVFMSNMSMDL-----QSHPSSCPCKDPCSC 194
 QY 179 KGSRCWSESESDOSLTRYACAGCA-RCKGRLPTDCHRCQCAAGCTGPKHSCLACLHF 237
 DB 195 PNGSCWGGEGENCOKLKIICAGQCSHRGSRSPDCCHNQCAGCTGPRESDCLVQCKF 254
 QY 238 NMSGICELHCPALVYNTDTFESMPNPEGRYTGASCVTACPYNYLSTDVGS 289
 DB 255 QDEATCKDTCPLMLNPTTYQMDVNEPKYSGATCVKRCPRNYVTVDHGS 306

RESULT 8
 Q9OX70 PRELIMINARY; PRT: 1209 AA.
 AC Q9OX70;
 ID Q9OX70;
 OX Q9OX70;

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Epidermal growth factor receptor.
 GN EGFR.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FISHER; TISSUE=LIVER;
 RX MEDLINE=9025888; PubMed=2342466;
 RA Pech L.A., Harris J., Raymond V.W., Blasband A.J., Lee D.C.,
 RA Earp H.S.;
 RT "A truncated, secreted form of the epidermal growth factor receptor is
 RT encoded by an alternatively spliced transcript in normal rat tissue."
 RL Mol. Cell. Biol. 10:2973-2982(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FISHER; TISSUE=LIVER;
 RA Pech L.A.;
 RL Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FISHER; TISSUE=LIVER;
 RA Guttridge K., Dawson T.L., Earp H.S.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; M37394; AAF14008.1; -.
 DR HSSP; P11362; JEGK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00261; Fy3; 3.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Receptor; Transferase; Tyrosine-protein kinase.
 KW SEQUENCE 1209 AA; 134891 MW; 96FEF7F6C1B773 CRC64;

Query Match 44.68; Score 708.5; DB 11; Length 1209;
 Best Local Similarity 47.48; Pred. No. 2.3e-58;
 Matches 137; Conservative 41; Mismatches 98; Indels 13; Gaps 4;

QY 3 QVCTGDMKRLPASPETHLMLRHLYOGCOVGNLETLPLPNASLFLDIOEVGY 62
 DB 29 KVCQGTNRRLTQLGFEDHFLSLQRMNCEVNLGNLEITVQNRNDLSFLKTIOEVGY 88
 QY 63 VLIANOVQVPLQRLRIVRGSTOLFEDNYALAVLDNGDPLNNTTPVYGASGGLREIQLR 122
 DB 89 FLIALNTERIPSEDLQIRGNALYENTYALALISN-----YGTNRTGLRELPMR 138
 QY 123 SLTEILKGVLIQNNPOLCYODTILMKDIFHKNNQALTLIDTNRK-RACHPCSPMKGS 181
 DB 139 NLQELIGAVRFNSNPILCNMTIOWRDI-VQDVFLSMMSMDVORHLILGCPKCPSCPNG 197
 QY 182 RCMWESSSDQSLRTVYAGGCA-RCKGPLPTDCHEOCAGCTGPKHSDCLACLHFHNS 240
 DB 198 SCWGEENCKLTKIICAOQCSRCRGRSPSDCHNCAGCTGCPRESDELCHREFDE 257
 QY 241 GICELHCPALVYNTDFESMPNEGRTYFGASCTACPYNYLSTVGS 289
 DB 258 ATCQDTQPLMLYNPTTYQMDVNPBGKYSFGATCVKCPRNYYVTDHGS 306

RESULT 9

Q9ESE0
 ID Q9ESE0 PRELIMINARY; PRT; 478 AA.
 AC Q9ESE0;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Epidermal growth factor receptor related protein.
 GN ERPP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=GASTRO-DUODENAL MUCOUS.
 RA Yu Y., Koshier J.A., Majumdar A.P.N.;
 RT "Cloning of a novel EGFR-related peptide: A putative negative
 RT regulator of EGFR."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF187818; AAG17037.2; -.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam; PF00757; Furin-like; 1.
 DR Pfam; PF01030; Recep_L_domain; 2.
 KW Receptor.
 KW SEQUENCE 478 AA; 53233 MW; CF873A8376C519E5 CRC64;

Query Match 44.18; Score 700.5; DB 11; Length 478;
 Best Local Similarity 46.68; Pred. No. 4.3e-58;
 Matches 136; Conservative 39; Mismatches 98; Indels 19; Gaps 4;

QY 3 QVCTGDMKRLPASPETHLMLRHLYOGCOVGNLETLPLPNASLFLDIOEVGY 62
 DB 29 KVCQGTNRRLTQLGFEDHFLSLQRMNCEVNLGNLEITVQNRNDLSFLKTIOEVGY 88
 QY 63 VLIANOVQVPLQRLRIVRGSTOLFEDNYALAVLDNGDPLNNTTPVYGASGGLREIQLR 122
 DB 89 FLIALNTERIPSEDLQIRGNALYENTYALALISN-----YGTNRTGLRELPMR 138
 QY 123 SLTEILKGVLIQNNPOLCYODTILMKDIFHKNNQALTLIDTNRKACHPCSPMC 178
 DB 139 NLQELIGAVRFNSNPILCNMTIOWRDIQVNPMSNMSML---GHPSSCRCPDSC 194
 QY 179 KGSRCWESSSDQSLRTVYAGGCA-RCKGPLPTDCHEOCAGCTGPKHSDCLACLHF 237
 DB 195 PNGSCWGEENCKLTKIICAOQCSRCRGRSPSDCHNCAGCTGPKHSDCLVCKF 254
 QY 238 NMSGICELHCPALVYNTDFESMPNEGRTYFGASCTACPYNYLSTVGS 289
 DB 255 QDEATCKDTQPLMLYNPTTYQMDVNPBGKYSFGATCVKCPRNYYVTDHGS 306

RESULT 10

Q9BUD7
 ID Q9BUD7 PRELIMINARY; PRT; 331 AA.
 AC Q9BUD7;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to v-erb-b2 avian erythroblastic leukemia viral oncogene
 DE homolog 3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC002706; AA022706.1; -.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.

QY 120 QLSLPEILGVLIOHPOLCYQDTILMKDI-FHKNNQALALIDITNRSRACHPCSPMC 178
 DB 105 ELPLALRIDLGSGVEFNNTLCHMKSSINMEIILAPOTSQMYTFNFSSPBPVPCPSPC 164
 QY 179 KGSRCMGESESDCOSLRTVCAGGCA--RCKGPLPTDCCHECOAAGCTGPKHSDCLACLH 236
 DB 165 EVG-CMGEAHNCORFESKLCSPQCSGRCFCGRPRECCHLFCAGCGCTGPTQSDCLACKN 223
 QY 237 FNHSGICELHCPALVYNTDTFESMPNPEGRTYFGASCYTACPYNYLSTDVGS 289
 DB 224 FYDDGVCKQCEPMPQIYNTPTNYFMEPRPDCKYAYGATCVRKCP-EHLKNDGA 275

RESULT 15

Q9W6F6 PRELIMINARY; PRT; 1137 AA.
 ID Q9W6F6
 AC Q9W6F6;
 DT 01-NOV-1999 (T-EMBLrel. 12, Created)
 DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Receptor tyrosine kinase (Fragment).
 GN ERBB4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HINDBRAIN;
 RX MEDLINE=99263203; PubMed=10328884;
 RA Dixon M., Lumsden A.;
 RT "Distribution of neuregulin-1 (nrg1) and erb4 transcripts in
 RL embryonic chick hindbrain.";
 DR EMBL: AF21963; AAD31764.1; -
 DR HSP: P11362; IFCK.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001368; TNFR-c6.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF01030; Recep_L_domain; 1.
 DR Pfam: PF02757; YLP; 2.
 DR PRINTS: PRO0109; TYRKINASE.
 DR PRODOM: PD000001; Euk_kinase; 1.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00219; TYRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 KM Kinase; Tyrosine-protein kinase.
 FT NON_TER
 SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87DC84F CRC64;

Query Match 25.6%; Score 407; DB 13; Length 1137;
 Best Local Similarity 48.3%; Pred. No. 8.8e-30;
 Matches 70; Conservative 18; Mismatches 55; Indels 2; Gaps 2;

QY 140 LCYQDTILMKDI-FHKNNQALALIDITNRSRACHPCSPMKGRMGESSESDCOSLRTVC 199
 DB 3 LCFADTTHMODIYRNPMASNFLLVPTNGSSGCRCHKSCGTG-RQWPTENHCQTLTKTVC 61
 QY 200 AGGC-ARCKGRLPTDCCHECOAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDTF 258
 DB 62 AEOCDGRCYGPVYDCGCHREAGCGSKPDTCFACMNFNDGACVYQCPOTFVYNTPTF 121
 QY 259 ESMNPNEGRTYFGASCYTACPYNYL 283

DB 122 QLEHNNAKTYGAFVYKCCPHNFV 146
 Search completed: April 28, 2003, 13:41:35
 Job time: 23.3311 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:39:54 ; Search time 8.45736 Seconds
(without alignments)
1005.423 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGYCTGDMKRLRASPET.....FGASCYACPYNTLSTVGS 289

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1587	100.0	419	4	US-09-630-155-2
2	1587	100.0	624	3	US-08-422-108-1
3	1587	100.0	624	4	US-08-422-734-1
4	1587	100.0	782	2	US-09-146-283-4
5	1587	100.0	782	3	US-08-579-823A-4
6	1587	100.0	782	4	US-09-344-195-4
7	1587	100.0	1235	1	US-08-467-083-68
8	1587	100.0	1235	1	US-08-414-417B-68
9	1587	100.0	1235	2	US-08-484-438-8
10	1587	100.0	1235	2	US-08-486-348A-68
11	1587	100.0	1235	2	US-08-625-101-2
12	1587	100.0	1235	2	US-08-468-545B-68
13	1587	100.0	1235	2	US-08-336-786-2
14	1587	100.0	1235	2	US-08-466-680B-68
15	1587	100.0	1235	1	US-08-336-708A-9
16	1587	100.0	1235	1	US-08-484-438-8
17	1587	100.0	1235	2	US-08-484-438-4
18	1587	100.0	1235	2	US-08-484-438-7
19	1587	100.0	1235	2	US-08-475-035-4
20	1587	100.0	1235	2	US-08-484-438-2
21	1587	100.0	1235	1	US-09-570-454-2
22	1587	100.0	1235	1	US-07-978-895-4
23	1587	100.0	1235	1	US-08-484-438-9
24	1587	100.0	1235	2	US-08-473-119-4
25	1587	100.0	1235	2	US-08-475-352-4
26	1587	100.0	1235	6	US-08-421-356-3
27	1587	100.0	1235	1	US-08-421-356-3

28	493	31.1	97	4	US-09-046-783-3	Sequence 3, Appl1
29	244.5	15.4	1382	2	US-08-737-715-2	Sequence 2, Appl1
30	244.5	15.4	1382	4	US-09-457-040B-7	Sequence 7, Appl1
31	232.5	14.7	516	3	US-08-746-559A-4	Sequence 4, Appl1
32	232.5	14.7	1367	2	US-08-249-687C-2	Sequence 2, Appl1
33	232.5	14.7	1367	2	US-08-625-819-2	Sequence 2, Appl1
34	232.5	14.7	1367	3	US-08-746-559A-2	Sequence 2, Appl1
35	232.5	14.7	1367	4	US-08-864-641B-18	Sequence 18, Appl1
36	231.5	14.6	486	3	US-08-746-559A-5	Sequence 5, Appl1
37	186	11.7	366	4	US-08-857-076-103	Sequence 103, App
38	185	11.7	1724	4	US-08-857-076-12	Sequence 12, Appl
39	171.5	10.8	383	4	US-08-857-076-105	Sequence 105, App
40	166.5	10.5	370	4	US-08-857-076-104	Sequence 104, App
41	132	8.3	381	4	US-08-857-076-106	Sequence 106, App
42	108.5	6.8	2211	4	US-09-738-884-1	Sequence 1, Appl1
43	103.5	6.5	769	2	US-08-789-078-1	Sequence 1, Appl1
44	103.5	6.5	769	2	US-08-752-633-1	Sequence 1, Appl1
45	103.5	6.5	769	2	US-08-476-062A-45	Sequence 45, Appl1

ALIGNMENTS

```
RESULT 1
US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130
GENERAL INFORMATION:
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTACONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2
Query Match
Best local similarity 100.0%; Score 1587; DB 4; Length 419;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 STGYCTGDMKRLRASPETHLDMRLHYOGGVQVGNLELYLPFNMSLFDIOEVO 60
DB 22 STGYCTGDMKRLRASPETHLDMRLHYOGGVQVGNLELYLPFNMSLFDIOEVO 81
QY 61 GYLLAHNOVROVPLQRLIVRGTOLEFDNVAVLAVLDNGDPLNNTTPTVGASPGGLRELO 120
```

DB 82 GYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQ 141
QY 121 LRSLEILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
DB 142 LRSLEILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201
QY 181 SRCGSESEDCQSLTRIVCAGGCAKRGPLPTDCHEQCAAGCTGPRHSDCLACIHFHNS 240
DB 202 SRCGSESEDCQSLTRIVCAGGCAKRGPLPTDCHEQCAAGCTGPRHSDCLACIHFHNS 261
QY 241 GICELHCPALVTYNTDFEESMPNEGRTFGASCVTACPNYISTDVG 289
DB 262 GICELHCPALVTYNTDFEESMPNEGRTFGASCVTACPNYISTDVG 310

RESULT 2
US-08-422-108-1
Sequence 1, Application US/08422108
Patent No. 6015567
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,108
FILING DATE: 14-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-APR-1993
APPLICATION DATA: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-108-1

Query Match 100.0%; Score 1587; DB 3; Length 624;
Best Local Similarity 100.0%; Pred. No. 6e-139;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STGVCTGTDKRLRLPASPEHLDMLRLHYGCGVVGNGLELTYLPTNASISFLDIOIEVO 60
DB 1 STGVCTGTDKRLRLPASPEHLDMLRLHYGCGVVGNGLELTYLPTNASISFLDIOIEVO 60

QY 61 GYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120
DB 61 GYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLDNGDPLNNTTPTVTGASPGGLRELQ 120
QY 121 LRSLEILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
DB 121 LRSLEILKGGVLIQRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
QY 181 SRCGSESEDCQSLTRIVCAGGCAKRGPLPTDCHEQCAAGCTGPRHSDCLACIHFHNS 240
DB 181 SRCGSESEDCQSLTRIVCAGGCAKRGPLPTDCHEQCAAGCTGPRHSDCLACIHFHNS 240
QY 241 GICELHCPALVTYNTDFEESMPNEGRTFGASCVTACPNYISTDVG 289
DB 241 GICELHCPALVTYNTDFEESMPNEGRTFGASCVTACPNYISTDVG 289

RESULT 3
US-08-422-734-1
Sequence 1, Application US/08422734
Patent No. 633169
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,734
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422108
FILING DATE: 14-Apr-1995
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
APPLICATION DATA: 07/354319
FILING DATE: 15-APR-1993
APPLICATION NUMBER: 08/048346
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-734-1

Query Match 100.0%; Score 1587; DB 4; Length 624;
Best Local Similarity 100.0%; Pred. No. 6e-139;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60
|
DB 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60
QY 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLRELO 120
|
DB 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLRELO 120
QY 121 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPCSPCKG 180
|
DB 121 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPCSPCKG 180
QY 181 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHFNHS 240
|
DB 181 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHFNHS 240
QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289
|
DB 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289

RESULT 4
US-09-146-283-4

Sequence 4, Application US/09146283
Patent No. 5976546
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-09-146-283-4

Query Match 100.0%; Score 1587; DB 2; Length 782;

Best Local Similarity 100.0%; Pred. No. 8, 1e-139;

Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60
|
DB 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60

DB 22 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 81
|
QY 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLRELO 120
|
DB 82 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLRELO 141
|
QY 121 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPCSPCKG 180
|
DB 142 LRSLEILKGGVLIQORNQOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPCSPCKG 201
|
QY 181 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHFNHS 240
|
DB 202 SRCWGESSEDCOSLTRFYCAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACLFHFNHS 261
|
QY 241 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 289
|
DB 262 GICELHCPALVTYNTDFESMPNEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 5
US-08-579-823A-4

Sequence 4, Application US/08579823A
Patent No. 6080409
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Ruegg, Curtis L.
APPLICANT: Wu, Hongyu
TITLE OF INVENTION: Immunostimulatory Composition and Method
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/579,823A
FILING DATE: 03-DEC-1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
US-08-579-823A-4

Query Match 100.0%; Score 1587; DB 3; Length 782;

Best Local Similarity 100.0%; Pred. No. 8, 1e-139;

Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 60
|
DB 22 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 81
|
QY 61 GYVLIAHNOVQVPLQRLRIYRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASPGGLRELO 120
|
DB 22 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVVOGNELETTYLPTNASLSFLDIOIEVQ 81

Db 82 GYVLIHNOVQVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 141
QY 121 LRSLETLKGVLIQRNPOLCYODTILMKDIFPKNNOLATLIDTNRSRACHCSPMCKG 180
Db 142 LRSLETLKGVLIQRNPOLCYODTILMKDIFPKNNOLATLIDTNRSRACHCSPMCKG 201
QY 181 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNS 240
Db 202 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNS 261
QY 241 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACAPNYLSTDVGS 289
Db 262 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACAPNYLSTDVGS 310

RESULT 6

US-09-344-195-4
Sequence 4, Application US/09344195
Patent No. 6210662
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
Ruegg, Curtis L.
Mu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0860
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 782 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein; Fig. 8
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-344-195-4

Query Match 100.0%; Score 1587; DB 4; Length 782;
Best Local Similarity 100.0%; Pred. No. 8.1e-139;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDMLKRLPASPEHLDMLRHLVGGCQVVGNNLELTPLPTNASISFLODIOEQV 60
Db 22 STQVCTGTDMLKRLPASPEHLDMLRHLVGGCQVVGNNLELTPLPTNASISFLODIOEQV 81

QY 61 GYVLIHNOVQVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 120
Db 82 GYVLIHNOVQVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 141
QY 121 LRSLETLKGVLIQRNPOLCYODTILMKDIFPKNNOLATLIDTNRSRACHCSPMCKG 180
Db 142 LRSLETLKGVLIQRNPOLCYODTILMKDIFPKNNOLATLIDTNRSRACHCSPMCKG 201
QY 181 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNS 240
Db 202 SRCWGSSEDCQSLTRTVACGAGCARCKGPLPTDCHEQCAAGCTGPKHSDCLACLFHNS 261
QY 241 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACAPNYLSTDVGS 289
Db 262 GICELHCPALVYNTDTFESMPNDEGRYTFGASCYACAPNYLSTDVGS 310

RESULT 7

US-08-467-083-68
Sequence 68, Application US/08467083
Patent No. 5726023
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,083
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/414,417
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEEDANBERY
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-467-083-68

Query Match 100.0%; Score 1587; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-138;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDMLKRLPASPEHLDMLRHLVGGCQVVGNNLELTPLPTNASISFLODIOEQV 60
Db 22 STQVCTGTDMLKRLPASPEHLDMLRHLVGGCQVVGNNLELTPLPTNASISFLODIOEQV 81
QY 61 GYVLIHNOVQVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 120
Db 82 GYVLIHNOVQVPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPTVGASGGLREIQ 141

Db 82 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLELQ 141
QY 121 LRSITELKGVLIQRRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
Db 142 LRSITELKGVLIQRRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201
QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGTGPRHSDCLACLFHNHS 240
Db 202 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGTGPRHSDCLACLFHNHS 261
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 289
Db 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 310

RESULT 10

US-08-486-348A-66
Sequence 68, Application US/08486348A
Patent No. 5846538
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-66

Query Match 100.0%; Score 1587; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-138;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGTDMKLRPLASPEHLDMLRHLVGGCOVGNLELTLYPTNASISFLDIDIOEQ 60
Db 22 STOVCTGTDMKLRPLASPEHLDMLRHLVGGCOVGNLELTLYPTNASISFLDIDIOEQ 81
QY 61 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLELQ 120
Db 82 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLELQ 141
QY 121 LRSITELKGVLIQRRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
Db 142 LRSITELKGVLIQRRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201

QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGTGPRHSDCLACLFHNHS 240
Db 202 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGTGPRHSDCLACLFHNHS 261
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 289
Db 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 310

RESULT 11

US-08-625-101-2
Sequence 2, Application US/08625101
Patent No. 5869445
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-Apr-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2

Query Match 100.0%; Score 1587; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-138;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STOVCTGTDMKLRPLASPEHLDMLRHLVGGCOVGNLELTLYPTNASISFLDIDIOEQ 60
Db 22 STOVCTGTDMKLRPLASPEHLDMLRHLVGGCOVGNLELTLYPTNASISFLDIDIOEQ 81
QY 61 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLELQ 120
Db 82 GYVLIANOVROVPLQRLRVRGTOLEEDNYALAVLDNGPPLNNTPTVTASPGGLELQ 141
QY 121 LRSITELKGVLIQRRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
Db 142 LRSITELKGVLIQRRPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201
QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGTGPRHSDCLACLFHNHS 240
Db 202 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGTGPRHSDCLACLFHNHS 261
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 289

Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 12

US-08-468-545B-68
; Sequence 68, Application US/08468545B
; Patent No. 5876712

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Disis, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,545B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C5

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

MOLECULE TYPE: linear

US-08-468-545B-68

Query Match

Best Local Similarity 100.0%; Score 1587; DB 2; Length 1255;

Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 STVCTGTDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNLSLFLDIOEVO 60

Db 22 STVCTGTDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNLSLFLDIOEVO 81

Db 61 GYVLIANQOVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 120

Db 82 GYVLIANQOVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 141

Db 121 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180

Db 142 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 201

Db 181 SRCWGESSEDCQSLRTVACAGCARKGRLPTDCHECCAGCTGPKHSDCLACLHFHNS 240

Db 202 SRCWGESSEDCQSLRTVACAGCARKGRLPTDCHECCAGCTGPKHSDCLACLHFHNS 261

Db 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289

Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 13

US-08-356-786-2
; Sequence 2, Application US/08356786

Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.

APPLICANT: Oppermann, Hermann

APPLICANT: Houston, L. L.

TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

TITLE OF INVENTION: Marker

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault

STREET: Exchange Place, 53 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,786

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7100

TELEFAX: (617) 248-7000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

MOLECULE TYPE: linear

US-08-356-786-2

Query Match

Best Local Similarity 100.0%; Score 1587; DB 2; Length 1255;

Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 STVCTGTDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNLSLFLDIOEVO 60

Db 22 STVCTGTDMKRLRPASPEHLDMLRHLVGGCQVVOGMLLETLYLPTNLSLFLDIOEVO 81

Db 61 GYVLIANQOVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 120

Db 82 GYVLIANQOVQVPLQRIYRGTOLEFDNYALAVLNGDPLNNTPTVTGASPGGLRELO 141

Db 121 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 180

Db 142 LRSLEILKGGVLIQIRNQLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPMKG 201

Db 181 SRCWGESSEDCQSLRTVACAGCARKGRLPTDCHECCAGCTGPKHSDCLACLHFHNS 240

Db 202 SRCWGESSEDCQSLRTVACAGCARKGRLPTDCHECCAGCTGPKHSDCLACLHFHNS 261

Db 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289

Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 14

US-08-466-680B-68
; Sequence 68, Application US/08466680B
; Patent No. 6073122

GENERAL INFORMATION:

```

APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,680B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ. ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-466-680B-68

Query Match          100.0%; Score 1587; DB 3; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.5e-138;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVCTGTDKRLRLPASPETHLMDLRHLRYGCGVQGNLELTLYLPTNASLSFLDIOIEVQ 60
DB 22 STVCTGTDKRLRLPASPETHLMDLRHLRYGCGVQGNLELTLYLPTNASLSFLDIOIEVQ 81
QY 61 GYVLIAHNOVROVPLORLRIYRGTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
DB 82 GYVLIAHNOVROVPLORLRIYRGTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141
QY 121 LRSLETILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPCKG 180
DB 142 LRSLETILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPCKG 201
QY 181 SRCMGESSEDCOSLTRRYCAGGACRCGKPLPTCCCHEQCAAGCTGPKHSDCLAFHNS 240
DB 202 SRCMGESSEDCOSLTRRYCAGGACRCGKPLPTCCCHEQCAAGCTGPKHSDCLAFHNS 261
QY 241 GICELHCPALVTYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 289
DB 262 GICELHCPALVTYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 15
US-08-336-708A-9
Sequence 9, Application US/08336708A
Patent No. 5521295
GENERAL INFORMATION:
APPLICANT: Pacificl, Robert E.
APPLICANT: Thomson, Arlen R.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: Hybrid Receptor Molecules
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.

```

```

STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,708A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy
REFERENCE/DOCKET NUMBER: A-241A
INFORMATION FOR SEQ. ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-708A-9

Query Match          45.1%; Score 716.5; DB 1; Length 644;
Best Local Similarity 46.5%; Pred. No. 2.9e-58;
Matches 134; Conservative 44; Mismatches 99; Indels 11; Gaps 2;

QY 3 QVCTGTDKRLRLPASPETHLMDLRHLRYGCGVQGNLELTLYLPTNASLSFLDIOIEVQ 62
DB 29 KVCQGSTNKLTLQGLTFEDHFLSLQRMFNCEVVLGNLEITYVORNDLSFLKTIQIEVAGY 88
QY 63 VLIATHNOVROVPLORLRIYRGTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 122
DB 89 VLIATHNOVROVPLORLRIYRGTOLEFDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 138
QY 123 SLTEILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPCKGSR 182
DB 139 NLGELIHGAVERFSNNALCNVESIQWRDIYSDPFLSNMSMDPNNHLSGCKCPGCPNGS 198
QY 183 CMGESSEDCOSLTRRYCAGGCA-RCKGRLPTDCCHEQCAAGCTGPKHSDCLAFHNSG 241
DB 199 CMGAGEBNCOKLTKIICAOQCSRCRCRCKSPSDCHNCAAGCTGPRESDCLVCRKFEDEA 258
QY 242 ICELHCPALVTYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 289
DB 259 TCKDTCPLMLNPTTYQMDVNEGKISFGATCVKCKCPRYNYVTDHGS 306

Search completed: April 28, 2003, 13:43:18
Job time : 15.4574 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:40:24 ; Search time 10.6321 Seconds

(Without alignments)
2178.074 Million cell updates/sec

Title: US-09-821-883-23

Perfect score: 1587

Sequence: 1 STGVCTGDMKRLRPASPER.....FGASCVTACPYNTLSTVGS 289

Scoring table: BLOSUM62

Gap: 10.0, Gapext 0.5

Search: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1587	100.0	289	10 US-09-821-883-23	Sequence 23, App1
2	1587	100.0	479	10 US-09-821-883-5	Sequence 5, App1
3	1587	100.0	555	10 US-09-821-883-1	Sequence 1, App1
4	1587	100.0	564	10 US-09-821-883-3	Sequence 3, App1
5	1587	100.0	645	10 US-09-821-161-1	Sequence 1, App1
6	1587	100.0	653	9 US-09-854-356-3	Sequence 3, App1
7	1587	100.0	690	10 US-09-821-883-2	Sequence 2, App1
8	1587	100.0	697	10 US-09-821-883-4	Sequence 4, App1
9	1587	100.0	712	9 US-09-854-356-7	Sequence 7, App1
10	1587	100.0	919	9 US-09-854-356-6	Sequence 6, App1
11	1587	100.0	1255	9 US-09-769-508-2	Sequence 2, App1
12	1587	100.0	1255	9 US-09-854-356-1	Sequence 1, App1
13	1587	100.0	1255	9 US-09-930-125-2	Sequence 2, App1
14	1587	100.0	1255	9 US-09-441-411-6	Sequence 6, App1
15	1587	100.0	1255	10 US-09-811-123-9	Sequence 3, App1
16	1587	100.0	1255	10 US-09-811-115-3	Sequence 3, App1
17	1353.5	85.3	654	9 US-09-854-356-8	Sequence 8, App1
18	1353.5	85.3	1256	9 US-09-854-356-2	Sequence 2, App1
19	1353.5	85.3	1260	9 US-09-870-759-118	Sequence 118, App

20	1352.5	85.2	1256	9 US-09-854-356-14	Sequence 14, App1
21	878	55.3	191	9 US-09-441-411-9	Sequence 9, App1
22	716.5	45.1	657	9 US-10-172-620-18	Sequence 18, App1
23	716.5	45.1	1210	10 US-09-725-433-2	Sequence 2, App1
24	716.5	45.1	1308	10 US-09-940-101-2	Sequence 4, App1
25	713.5	45.0	615	10 US-09-940-101-4	Sequence 2, App1
26	700.5	44.1	478	10 US-09-867-521-2	Sequence 16, App1
27	686	43.2	1342	9 US-10-172-620-16	Sequence 16, App1
28	249	15.7	120	9 US-10-172-620-17	Sequence 17, App1
29	232.5	14.7	1367	9 US-09-870-759-120	Sequence 120, App
30	186	11.7	366	10 US-09-205-658-103	Sequence 103, App
31	186	11.7	366	10 US-09-844-353A-103	Sequence 103, App
32	185	11.7	1724	10 US-09-205-658-12	Sequence 12, App1
33	185	11.7	1724	10 US-09-844-353A-12	Sequence 12, App1
34	171.5	10.8	383	10 US-09-205-658-105	Sequence 105, App
35	171.5	10.8	383	10 US-09-844-353A-105	Sequence 105, App
36	166.5	10.5	370	10 US-09-205-658-104	Sequence 104, App
37	166.5	10.5	370	10 US-09-844-353A-104	Sequence 104, App
38	132	8.3	381	10 US-09-205-658-106	Sequence 106, App
39	132	8.3	381	10 US-09-844-353A-106	Sequence 106, App
40	117	7.4	2290	9 US-10-123-155-123	Sequence 323, App
41	116.5	7.3	68	10 US-09-466-320-11	Sequence 11, App1
42	116	7.3	22	10 US-09-466-320-19	Sequence 19, App1
43	115.5	7.3	1875	9 US-10-123-155-359	Sequence 359, App
44	114.5	7.2	1281	9 US-10-123-155-509	Sequence 509, App
45	114	7.2	840	9 US-10-123-155-227	Sequence 227, App

ALIGNMENTS

RESULT 1
US-09-821-883-23
Sequence 23, Application US/09821863
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
TITLE OF INVENTION: Gradis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT FILING DATE: 2001-03-30
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 23
LENGTH: 289
TYPE: PRT
ORGANISM: Homo sapiens
US-09-821-883-23
Query Match 100.0%; Score 1587; DB 10; Length 289;
Best Local Similarity 100.0%; Pred. No. 6.2e-120;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 STGVCTGDMKRLRPASPERHLMRLHYGCGVYQWLELTLYPTNASTSFLDIOEVQ 60
DB 1 STGVCTGDMKRLRPASPERHLMRLHYGCGVYQWLELTLYPTNASTSFLDIOEVQ 60
QY 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLNDGDPPLNNTPTVTSASPGGLRELQ 120
DB 61 GYVLIANQVQVPLQRLRIYRGTOLEFDNYALAVLNDGDPPLNNTPTVTSASPGGLRELQ 120
QY 121 LRSLEILKGVGLIQRNPOLCYODTILMKDIFHKNNQALATLIDTNRACHPSCPMKG 180
DB 121 LRSLEILKGVGLIQRNPOLCYODTILMKDIFHKNNQALATLIDTNRACHPSCPMKG 180
QY 181 SRCWESSEDCOSLTRIVCAGGACRCARCPPLTDCHEOCAGCTGPKHSXCLACHFNHS 240
DB 181 SRCWESSEDCOSLTRIVCAGGACRCARCPPLTDCHEOCAGCTGPKHSXCLACHFNHS 240

QY 241 GICELHCPALVTYNTDTFESMPNPEGRTGASCVTACPNYLISTDVG 289
DB 241 GICELHCPALVTYNTDTFESMPNPEGRTGASCVTACPNYLISTDVG 289

RESULT 2

US-09-821-883-5
Sequence 5, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 479
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER300*-rgm-CSF construct
US-09-821-883-5

Query Match 100.0%; Score 1587; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.1e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKLRIPASPEHMLRLHYGCGVGNLELTLPYNASLSFLODIOEVQ 60
DB 41 STQVCTGDMKLRIPASPEHMLRLHYGCGVGNLELTLPYNASLSFLODIOEVQ 100
QY 61 GYVLIANOVROVPLQRLIRVGTQLEFDNYALAVDNGDPLNNTTPVTGASGGLREIQ 120
DB 101 GYVLIANOVROVPLQRLIRVGTQLEFDNYALAVDNGDPLNNTTPVTGASGGLREIQ 160
QY 121 LRSLETLKGVLIORNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCSPCKG 180
DB 161 LRSLETLKGVLIORNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCSPCKG 220
QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 240
DB 221 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 280
QY 241 GICELHCPALVTYNTDTFESMPNPEGRTGASCVTACPNYLISTDVG 289
DB 281 GICELHCPALVTYNTDTFESMPNPEGRTGASCVTACPNYLISTDVG 329

RESULT 3

US-09-821-883-1
Sequence 1, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1
LENGTH: 555
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 100.0%; Score 1587; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.3e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKLRIPASPEHMLRLHYGCGVGNLELTLPYNASLSFLODIOEVQ 60
DB 41 STQVCTGDMKLRIPASPEHMLRLHYGCGVGNLELTLPYNASLSFLODIOEVQ 100
QY 61 GYVLIANOVROVPLQRLIRVGTQLEFDNYALAVDNGDPLNNTTPVTGASGGLREIQ 120
DB 101 GYVLIANOVROVPLQRLIRVGTQLEFDNYALAVDNGDPLNNTTPVTGASGGLREIQ 160
QY 121 LRSLETLKGVLIORNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCSPCKG 180
DB 161 LRSLETLKGVLIORNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCSPCKG 220
QY 181 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 240
DB 221 SRCWGESSEDCQSLTRTVACGAGCARCKGPLPTDCHBOCAAGCTGPKHSDCLACLFHNHS 280
QY 241 GICELHCPALVTYNTDTFESMPNPEGRTGASCVTACPNYLISTDVG 289
DB 281 GICELHCPALVTYNTDTFESMPNPEGRTGASCVTACPNYLISTDVG 329

RESULT 4

US-09-821-883-3
Sequence 3, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 564
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 100.0%; Score 1587; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.3e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGDMKLRIPASPEHMLRLHYGCGVGNLELTLPYNASLSFLODIOEVQ 60
DB 41 STQVCTGDMKLRIPASPEHMLRLHYGCGVGNLELTLPYNASLSFLODIOEVQ 100
QY 61 GYVLIANOVROVPLQRLIRVGTQLEFDNYALAVDNGDPLNNTTPVTGASGGLREIQ 120
DB 101 GYVLIANOVROVPLQRLIRVGTQLEFDNYALAVDNGDPLNNTTPVTGASGGLREIQ 160
QY 121 LRSLETLKGVLIORNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCSPCKG 180
DB 161 LRSLETLKGVLIORNPOLCYODTILMKDIFKNNOLATLIDTNRSRACHPCSPCKG 220

```

QY 181 SRCWSESSEDCOSLTRFYCAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 240
    |||||||
Db 221 SRCWSESSEDCOSLTRFYCAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 280
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289
    |||||||
Db 281 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 329

```

RESULT 5

```

US-09-921-161-1
: Sequence 1, Application US/09921161
: Patent No. US20020090662A1
: GENERAL INFORMATION:
: APPLICANT: Ralph, Peter
: TITLE OF INVENTION: ANALYTICAL METHOD
: FILE REFERENCE: GENENT.066A
: CURRENT APPLICATION NUMBER: US/09/921,161
: PRIOR FILING DATE: 2001-08-01
: PRIOR APPLICATION NUMBER: 60/225,433
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 645
: TYPE: PRF
: ORGANISM: Homo sapiens
US-09-921-161-1

```

```

Query Match          100.0%; Score 1587; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 1,6e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STQVCTGDMKRLRPASPEHLDMLRHLYGCGVQVGNLELTYPNTASLSFLDIOEYQ 60
    |||||||
Db 22 STQVCTGDMKRLRPASPEHLDMLRHLYGCGVQVGNLELTYPNTASLSFLDIOEYQ 81
QY 61 GYVLIHNOVQVPLQRLIRYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120
    |||||||
Db 82 GYVLIHNOVQVPLQRLIRYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNRSRACHPCSPMKKG 180
    |||||||
Db 142 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNRSRACHPCSPMKKG 201
QY 181 SRCWSESSEDCOSLTRFYCAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 240
    |||||||
Db 202 SRCWSESSEDCOSLTRFYCAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289
    |||||||
Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

```

RESULT 6

```

US-09-854-356-3
: Sequence 3, Application US/09854356
: Patent No. US2002017367A1
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Cheyssen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-26
: PRIOR APPLICATION NUMBER: US 60/117,976
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 26

```

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 653
: TYPE: PRF
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: extracellular domain (ECD) of human HER-2/neu
US-09-854-356-3

```

```

Query Match          100.0%; Score 1587; DB 9; Length 653;
Best Local Similarity 100.0%; Pred. No. 1,6e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STQVCTGDMKRLRPASPEHLDMLRHLYGCGVQVGNLELTYPNTASLSFLDIOEYQ 60
    |||||||
Db 22 STQVCTGDMKRLRPASPEHLDMLRHLYGCGVQVGNLELTYPNTASLSFLDIOEYQ 81
QY 61 GYVLIHNOVQVPLQRLIRYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120
    |||||||
Db 82 GYVLIHNOVQVPLQRLIRYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNRSRACHPCSPMKKG 180
    |||||||
Db 142 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNRSRACHPCSPMKKG 201
QY 181 SRCWSESSEDCOSLTRFYCAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 240
    |||||||
Db 202 SRCWSESSEDCOSLTRFYCAGCARCKGRLPTDCHECCAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 289
    |||||||
Db 262 GICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

```

RESULT 7

```

US-09-821-883-2
: Sequence 2, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Vidovic, Damar
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 690
: TYPE: PRF
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500-hgm-CSF construct
US-09-821-883-2

```

```

Query Match          100.0%; Score 1587; DB 10; Length 690;
Best Local Similarity 100.0%; Pred. No. 1,7e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 STQVCTGDMKRLRPASPEHLDMLRHLYGCGVQVGNLELTYPNTASLSFLDIOEYQ 60
    |||||||
Db 41 STQVCTGDMKRLRPASPEHLDMLRHLYGCGVQVGNLELTYPNTASLSFLDIOEYQ 100
QY 61 GYVLIHNOVQVPLQRLIRYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 120
    |||||||
Db 101 GYVLIHNOVQVPLQRLIRYRGTOLEFDNYALAVLDNGDPLNNTPTVYGASPGGLRELO 160
QY 121 LRSLEILKGGVLIQIRNPOLCYODTIIMKDIFFHKNQALATLIDTNRSRACHPCSPMKKG 180
    |||||||

```

Db 161 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 220
QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 240
Db 221 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 280
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 289
Db 281 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 329

RESULT 8
US-09-821-883-4
Sequence 4, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Retner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE OF INVENTION: Cell-Based Immunotherapy
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883.
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 697
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500*-igm-CSF construct
US-09-821-883-4

Query Match 100.0%; Score 1587; DB 10; Length 697;
Best Local Similarity 100.0%; Pred. No. 1.7e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDMKRLRPA SPETHLDMRLHYGCGVGNELTYLPTNASTSFLQDIOEVQ 60
Db 41 STQVCTGTDMKRLRPA SPETHLDMRLHYGCGVGNELTYLPTNASTSFLQDIOEVQ 100
QY 61 GYVLIHNOVROYPLQRLIRVGTOLFEEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
Db 101 GYVLIHNOVROYPLQRLIRVGTOLFEEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 160
QY 121 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
Db 161 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 220
QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 240
Db 221 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 280
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 289
Db 281 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 329

RESULT 9
US-09-854-356-7
Sequence 7, Application US/09854356
Patent No. US20020177567A1
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheysen, Dirk
APPLICANT: Corixa Corporation
APPLICANT: SmithKline Beecham Biologicals S. A.
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854,356

QY 1 STQVCTGTDMKRLRPA SPETHLDMRLHYGCGVGNELTYLPTNASTSFLQDIOEVQ 60
Db 22 STQVCTGTDMKRLRPA SPETHLDMRLHYGCGVGNELTYLPTNASTSFLQDIOEVQ 81
QY 61 GYVLIHNOVROYPLQRLIRVGTOLFEEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
Db 82 GYVLIHNOVROYPLQRLIRVGTOLFEEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141
QY 121 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
Db 142 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201
QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 240
Db 202 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 261
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 289
Db 262 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 310

Query Match 100.0%; Score 1587; DB 9; Length 712;
Best Local Similarity 100.0%; Pred. No. 1.8e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STQVCTGTDMKRLRPA SPETHLDMRLHYGCGVGNELTYLPTNASTSFLQDIOEVQ 60
Db 22 STQVCTGTDMKRLRPA SPETHLDMRLHYGCGVGNELTYLPTNASTSFLQDIOEVQ 81
QY 61 GYVLIHNOVROYPLQRLIRVGTOLFEEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
Db 82 GYVLIHNOVROYPLQRLIRVGTOLFEEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141
QY 121 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 180
Db 142 LRSLETLKGGVLIORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPCKG 201
QY 181 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 240
Db 202 SRCWGESSEDCQSLTRTVACGAGCARGPLPTDCHEBOCAAGCTGPKHSDCLACLFHNHS 261
QY 241 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 289
Db 262 GICELHCPALVYNTDTFESMPNPEGRTYFGASCVCATPCPNYLISTDVG 310

Query Match 100.0%; Score 1587; DB 9; Length 919;
Best Local Similarity 100.0%; Pred. No. 2.4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 60
D 22 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 81
QY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 120
D 82 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 180
D 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 201
QY 181 SRCWESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
D 202 SRCWESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 289
D 262 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 310
```

RESULT 11

```
US-09-769-508-2
Sequence 2, Application US/09769508
Patent No. US2002015527A1
GENERAL INFORMATION:
APPLICANT: STUART, SUSAN G.
APPLICANT: MONAHAN, JOHN J.
APPLICANT: LANGTON, BEATRICE CLAUDIA
APPLICANT: HANCOCK, MIRIAM E.C.
APPLICANT: CHAO, LORRINE A.
APPLICANT: BLUFORD, PETER
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
FILE REFERENCE: BBIO-111-C1
CURRENT APPLICATION NUMBER: US/09/769, 508
CURRENT FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-769-508-2
```

```
Query Match 100.0%; Score 1587; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 60
D 22 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 81
QY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 120
D 82 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 180
D 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 201
QY 181 SRCWESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
D 202 SRCWESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 289
D 262 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 310
```

RESULT 12

```
US-09-854-356-1
Sequence 1, Application US/09854356
Patent No. US2002017567A1
```

```
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Cheyssen, Dirk
APPLICANT: Corixa Corporation
APPLICANT: SmithKline Beecham Biologicals S. A.
TITLE OF INVENTION: HER-2/neu Fusion Proteins
FILE REFERENCE: 014058-009810PC
CURRENT APPLICATION NUMBER: US/09/854, 356
CURRENT FILING DATE: 2001-05-09
PRIORITY APPLICATION NUMBER: US 09/493,480
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: US 60/117,976
PRIORITY FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human HER-2/neu protein
NAME/KEY: DOMAIN
LOCATION: (1)..(653)
OTHER INFORMATION: extracellular domain (ECD)
NAME/KEY: DOMAIN
LOCATION: (676)..(1255)
OTHER INFORMATION: intracellular domain (ICD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1255)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (990)..(1048)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
US-09-854-356-1
```

```
Query Match 100.0%; Score 1587; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 60
D 22 STOVCTGDMKRLRPASPEHLDMLRHLHYGCGVQVQGNLELTYLPTNASLSFLDIOEQV 81
QY 61 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 120
D 82 GYVLIAHQVQVPLQRLRIYRGTOLEFDNYALAVLDNGDPLNNTPTVTGASPGGLRELO 141
QY 121 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 180
D 142 LRSLEILKGGVLIQORNQOLCYODTILMKDIFHKNNQALATLIDTNRBRACHPCSPMCKG 201
QY 181 SRCWESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 240
D 202 SRCWESSEDCQSLRTVYACAGCARGKPLPTDCHEQCAAGCTGPKHSDCLACHFNHS 261
QY 241 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 289
D 262 GICELHCPALVTYNTDFESMPNPEGRTYFGASCTYACPYNYLSTDVGS 310
```

RESULT 13

```
US-09-930-125-2
Sequence 2, Application US/09930125
Publication No. US20020193329A1
```

```
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vedvick, Thomas S.
```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-125-2
```

```

Query Match          100.0%; Score 1587; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCVVGNLELTYLPTNASLSFLDIOIEVQ 60
DB 22 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCVVGNLELTYLPTNASLSFLDIOIEVQ 81
OY 61 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
DB 82 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141
OY 121 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKCG 180
DB 142 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKCG 201
OY 181 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPHSDCLACLHFNHS 240
DB 202 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPHSDCLACLHFNHS 261
OY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 289
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 310
```

```

RESULT 14
US-09-441-411-6
; Sequence 6, Application US/09441411
; Publication No. US20030008342A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Scholler, Nathalie B.
; APPLICANT: Disis, Mary L.
; APPLICANT: Helstrom, Ingegerd
; APPLICANT: Helstrom, Karl Erik
; TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
; FILE REFERENCE: 730033.409
; CURRENT APPLICATION NUMBER: US/09/441.411
; CURRENT FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-441-411-6
```

```

Query Match          100.0%; Score 1587; DB 9; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCVVGNLELTYLPTNASLSFLDIOIEVQ 60
DB 22 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCVVGNLELTYLPTNASLSFLDIOIEVQ 81
OY 61 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
DB 82 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141
OY 121 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKCG 180
```

```

DB 142 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKCG 201
OY 181 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPHSDCLACLHFNHS 240
DB 202 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPHSDCLACLHFNHS 261
OY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 289
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 310
```

```

RESULT 15
US-09-811-123-9
; Sequence 9, Application US/09811123
; Patent No. US20020001587A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Sharon Erickson
; APPLICANT: Ralph Schwall
; APPLICANT: Mark Sliwowski
; TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBB
; FILE REFERENCE: GENENT.07342
; CURRENT APPLICATION NUMBER: US/09/811.123
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/238.327
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 09/602.530
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-123-9
```

```

Query Match          100.0%; Score 1587; DB 10; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.4e-119;
Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

OY 1 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCVVGNLELTYLPTNASLSFLDIOIEVQ 60
DB 22 STOVCTGTDKMLRLPASPEHLDMLRHLVGGCVVGNLELTYLPTNASLSFLDIOIEVQ 81
OY 61 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 120
DB 82 GYVLIANNOVROYPLQRLRIVRGTOLEFEDNYALAVLDNGDPLNNTTPVTGASPGGLREIQ 141
OY 121 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKCG 180
DB 142 LRSITELKGGVLIQRRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPMKCG 201
OY 181 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPHSDCLACLHFNHS 240
DB 202 SRCWGESSEDCOSLTRIVCAGGACRCKGRLPTDCHEQCAAGCTGPHSDCLACLHFNHS 261
OY 241 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 289
DB 262 GICELHCPALVYNTDTFESMPNPEGRTYTGASCVTACPYNYLSTDVGS 310
```

```

Search completed: April 28, 2003, 13:44:07
Job time : 14.6321 secs
```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:43:24 ; Search time 22 seconds
(without alignments)
3015.125 million cell updates/sec

Title: US-09-821-883-2

Perfect score: 690
Sequence: 1 MRAAPLLARASLSIGFLF.....EPVQCAPPPPAHHHHHH 690

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	292	42.3	1255	1 A24571	protein-tyrosine k
2	127	18.4	144	1 FOHUCM	granulocyte-macrop
3	67	9.7	1234	2 I48161	p-185 precursor
4	52	7.5	1260	1 TVRTNU	protein-tyrosine k
5	35	5.1	386	1 JH0610	acid phosphatase (
6	22	3.2	144	1 A61632	granulocyte-macrop
7	22	3.2	144	2 JH0469	granulocyte-macrop
8	11	1.6	144	2 A44936	granulocyte-macrop
9	10	1.4	143	1 FOHUCM	granulocyte-macrop
10	9	1.3	527	2 A42032	epidermal growth f
11	9	1.3	644	2 A36325	epidermal growth f
12	9	1.3	1210	1 COHUE	epidermal growth f
13	9	1.3	1210	2 A53183	epidermal growth f
14	9	1.3	1223	1 TVCHRV	epidermal growth f
15	8	1.2	127	2 I46269	granulocyte-macrop
16	8	1.2	265	2 A46534	ubiquitin-cytochro
17	8	1.2	268	2 A56446	ig heavy chain v r
18	8	1.2	325	2 D70666	probable mod prot
19	8	1.2	341	2 B83298	conserved hypotet
20	8	1.2	348	2 D95067	phenylalanyl-trna
21	8	1.2	354	2 F87388	hyd family secret
22	8	1.2	355	2 T44850	molybdopterin co-f
23	8	1.2	374	2 I39637	molybdopterin co-f
24	8	1.2	375	2 C97935	phenylalanine-trna
25	8	1.2	463	2 H70922	hypothetical prote
26	8	1.2	517	2 T31463	probable diaphyt
27	8	1.2	539	2 AH0508	sensor kinase cita
28	8	1.2	547	2 S70538	signal-transducing
29	8	1.2	583	2 D90052	hypothetical prote

30	8	1.2	610	2 T32327	hypothetical prote
31	8	1.2	715	2 S73637	ATP-dependent pr
32	8	1.2	825	1 EDBEXD	immediate-early pr
33	8	1.2	1851	2 T13980	calcium channel al
34	7	1.0	62	2 F90177	hypothetical prote
35	7	1.0	65	2 T29066	hypothetical prote
36	7	1.0	68	2 E69489	LSU ribosomal prot
37	7	1.0	82	2 T30135	hypothetical prote
38	7	1.0	89	2 A10221	flagellar biosynth
39	7	1.0	98	2 E90606	hypothetical prote
40	7	1.0	104	1 G69482	hypothetical prote
41	7	1.0	111	2 C35826	hypothetical 13k p
42	7	1.0	114	2 E72600	hypothetical prote
43	7	1.0	120	2 T35946	probable regulator
44	7	1.0	125	1 TCCHRP	calcitonin gene-re
45	7	1.0	136	1 TCON	calcitonin 1 precu

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein e
C:Species: Homo sapiens (man)
C>Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999
C:Accession: A24571; A23491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:931197; PIDN:CAA27060.1; PID:931198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A23491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M1767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Coonsens, L.; Yang-Feng T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALU', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M1730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Molecule type: DNA
A:Status: translated from GB/EMBL/DBJ
A:Residues: 832-909 <REX>
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAU>

A:Cross-references: GB:M16792; NID:g183983; P1DN:AAA58637.1; PID:g553332
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-17q21.1
 A:Introns: 25/1; 75/3; 147/1; 883/3
 A:Note: The list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 lase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-155/Product: protein-tyrosine kinase erbB2 #status predicted <MAY>
 F:22-653/Domain: extracellular #status predicted <EXT>
 F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
 F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
 F:654-675/Domain: transmembrane #status predicted <TM>
 F:676-1255/Domain: intracellular #status predicted <INT>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 F:68,124,187,259,530,571,629/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:753/Active site: Lys #status predicted
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 42.3%; Score 292; DB 1; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 6, 9e-281;
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAATGCTGCTGDMKLRPASPETHLDMRLHLYGCCVYVCCLELTPTNALSFLDDIO 97
 DB 19 GAAATGCTGCTGDMKLRPASPETHLDMRLHLYGCCVYVCCLELTPTNALSFLDDIO 78
 OY 98 EVGGVLIANNOVROVPLQRLRYGTQLFEDNVALAVLNDGDLNNTPTPTGSPGCLR 157
 DB 79 EVGGVLIANNOVROVPLQRLRYGTQLFEDNVALAVLNDGDLNNTPTPTGSPGCLR 138
 OY 158 ELQRLSTLEIKGGVLIQRRNPOLCYQDTILMKDIFHRNNOLATLIDTNRSRACPCSPM 217
 DB 139 ELQRLSTLEIKGGVLIQRRNPOLCYQDTILMKDIFHRNNOLATLIDTNRSRACPCSPM 198
 OY 218 CKGRCKGSESEDCOSLTRVYACAGCARGPLPTDCCHQCAGCGPRHSDCLALHF 277
 DB 199 CKGRCKGSESEDCOSLTRVYACAGCARGPLPTDCCHQCAGCGPRHSDCLALHF 258
 OY 278 NMSGICLHCPALVTYNTDFESMPNREGRTFGASCVTACPPYNYLSTDVGS 329
 DB 259 NMSGICLHCPALVTYNTDFESMPNREGRTFGASCVTACPPYNYLSTDVGS 310

RESULT 2

FOHUGM

granulocyte-macrophage colony-stimulating factor precursor (validated) - human

N:Alternate names: colony-stimulating factor 2; GM-CSF

C:Species: Homo sapiens (man)

C:Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 08-Dec-2000

C:Accession: C24636; 159065; A25169; A44175; JCI1090

R:Myakeke, S.; Otsuka, T.; Yokota, T.; Lee, F.; Aral, K.

EMBO J. 4, 2561-2568, 1985

A:Title: Structure of the chromosomal gene for granulocyte-macrophage colony stimulating

A:Reference number: A91015; M0ID:86030234; PMID:3876930

A:Accession: C24636

A:Molecule type: DNA

A:Residues: 1-144 <MY>

A:Cross-references: EMBL:X03021; NID:g31858; P1DN:CAA26822.1; PID:g31859

R:Kaushtansky, K.; O'Hara, P.J.; Berkner, K.; Segal, G.M.; Hagen, F.S.; Adamson, J.W.

Proc. Natl. Acad. Sci. U.S.A. 83, 3101-3105, 1986

A:Title: Genomic cloning, characterization, and multilignage growth-promoting activity

C:Accession: 159065; M0ID:86205844; PMID:3486413

A:Reference number: 159065

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-144 <RS>
 A:Cross-references: GB:M13207; NID:g181147; P1DN:AAA98768.1; PID:g181148
 R:Cantrrell, M.A.; Anderson, D.; Ceretti, D.P.; Price, V.; McKereghan, K.; Tushnetsk,
 Proc. Natl. Acad. Sci. U.S.A. 82, 6250-6254, 1985
 A:Title: Cloning, sequence, and expression of a human granulocyte/macrophage colony-s
 A:Reference number: A25169; M0ID:85298329; PMID:3898082
 A:Accession: A25169
 A:Molecule type: mRNA
 A:Residues: 1-144 <CAN>
 A:Cross-references: GB:M1734; NID:g181149; P1DN:AAA52122.1; PID:g181150
 R:Lee, F.; Yokota, T.; Otsuka, T.; Gemmell, L.; Larson, N.; Luh, J.; Aral, K.; Rennic
 Proc. Natl. Acad. Sci. U.S.A. 82, 4360-4364, 1985
 A:Title: Isolation of cDNA for a human granulocyte-macrophage colony-stimulating fact
 A:Reference number: A01853; M0ID:85242684; PMID:3925454
 A:Accession: A01853
 A:Molecule type: mRNA
 A:Residues: 1-144 <LEE>
 A:Cross-references: GB:M11720; NID:g183363; P1DN:AAA52578.1; PID:g183364
 R:Wong, G.G.; Witke, J.S.; Temple, P.A.; Wilkens, K.M.; Leary, A.C.; Luxenberg, D.P.;
 A:Clark, S.C.
 Science 228, 810-815, 1985
 A:Title: Human GM-CSF: molecular cloning of the complementary DNA and purification of
 A:Reference number: A44175; M0ID:85218749; PMID:3923623
 A:Accession: A44175
 A:Molecule type: mRNA
 A:Residues: 1-116, 'T', 118-144 <MON>
 A:Cross-references: GB:M10663; NID:g181145; P1DN:AAA52121.1; PID:g181146
 A:Note: parts of this sequence, including the amino end of the mature protein, were c
 R:Wen, D.Y.; Huang, B.R.; Cai, L.W.; Si, J.Y.
 Acta Biochim. Biophys. Sin. 25, 651-655, 1993
 A:Title: Amplification of human granulocyte-macrophage colony-stimulating factor cDNA
 A:Reference number: JCI1090
 A:Accession: JCI1090
 A:Molecule type: protein
 A:Residues: 18-21, 'C', 23-96, 'L', 98-144 <WEN>
 C:Genetics:
 A:Gene: GDB:CSF2
 A:Cross-references: GDB:119812; OMIM:138960
 A:Map position: 5q23.2-5q31.1
 A:Introns: 53/3; 67/3; 109/3
 C:Function:
 A:Description: stimulates the differentiation and proliferation of hematopoietic prog
 C:Superfamily: granulocyte-macrophage colony-stimulating factor
 C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-144/Product: granulocyte-macrophage colony-stimulating factor #status experiment
 F:44,54/Binding site: carboxydrate (Asn) (covalent) #status predicted

Query Match 18.4%; Score 127; DB 1; Length 144;
 Best Local Similarity 100.0%; Pred. No. 6, 5e-118;
 Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 549 APARSPSTQPMHVAIOEARLLNLSDTAENMEVVEISEMPDLOEPTCLOTRLE 608
 DB 18 APARSPSTQPMHVAIOEARLLNLSDTAENMEVVEISEMPDLOEPTCLOTRLE 77
 OY 609 LYKQGLGSLTKLKGPLTMASHYKCHCPPTPETSATQITTFESFENKDFLVLVFD 668
 DB 78 LYKQGLGSLTKLKGPLTMASHYKCHCPPTPETSATQITTFESFENKDFLVLVFD 137
 OY 669 CMEPVQD 675
 DB 138 CMEPVQD 144

RESULT 3

I48161

P-185 precursor - golden hamster

C:Species: Mesocricetus auratus (golden hamster)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999

C:Accession: I48161

R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagaoka, M.; Aral, M.; Yamazaki, Y.; Ishika

Gene 140, 251-255, 1994
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
 A:Reference number: 148161; MUID:94193007; PMID:7908275
 A:Accession: 148161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1254 <RES>
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 9.7%; Score 67; DB 2; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 1.6e-57;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 250 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRT 309
 |||
 Db 231 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESMPNEGRT 290

Oy 310 FGASCVT 316
 |||
 Db 291 FGASCVT 297

RESULT 4
 TVRTNU
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C:Accession: A24562; A61204
 R:Baragman, C.I.; Hung, M.C.; Weinberg, R.A.
 Nature 319, 226-230, 1986
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
 A:Reference number: A24562; MUID:86118662; PMID:3945311
 A:Accession: A24562
 A:Molecule type: mRNA
 A:Residues: 1-1260 <BAR>
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746
 R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen, C.
 Carcinogenesis 12, 1975-1978, 1991
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no p
 2-thiazolyl]formamide or N-methyl-N-nitrosourea.
 A:Reference number: A61204; MUID:92035293; PMID:1682063
 A:Accession: A61204
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 637-663, 'V', 665-702 <MAS>
 A:Note: authors translated the codon GCA for residue 25 as Val
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>
 F:656-680/Domain: transmembrane #status predicted <TMN>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:711,191,263,535,576,634/Binding site: carbonylrate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:756/Active site: lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 7.5%; Score 52; DB 1; Length 1260;
 Best Local Similarity 100.0%; Pred. No. 1.3e-42;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 250 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESM 301
 |||
 Db 235 LPTDCHEOCAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDFESM 286

RESULT 5
 JH0610
 acid phosphatase (EC 3.1.3.2) ACP precursor [validated] - human
 N:Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphohydr
 C:Species: Homo sapiens (man)
 C:Date: 17-Aug-1992 #sequence_revision 01-Dec-1995 #text_change 08-Dec-2000
 C:Accession: JH0610; J50693; A38608; S01331; A32419; S11147; S38863; S41251; S17042;
 R:Shartel, F.S.; Li, S.S.L.
 Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992
 A:Title: Structure of human prostatic acid phosphatase gene.
 A:Reference number: JH0610; MUID:92272747; PMID:1375464
 A:Accession: JH0610
 A:Molecule type: DNA
 A:Residues: 1-386 <SHA>
 A:Cross-references: GB:M97580; GB:M97581; GB:M97582; GB:M97583; GB:M97584; GB:M97585;
 A:Accession: J50693
 A:Molecule type: mRNA
 A:Residues: 1-386 <SH3>
 A:Cross-references: GB:M97589; NID:9189611; PIDN:AAA60021.1; PID:9189613
 R:Van Etten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.
 J. Biol. Chem. 266, 2313-2319, 1991
 A:Title: Covalent structure, disulfide bonding, and identification of reactive surfac
 A:Reference number: A38608; MUID:91115848; PMID:1989985
 A:Accession: A38608
 A:Molecule type: mRNA
 A:Residues: 1-386 <VAN>
 A:Cross-references: GB:M24840; NID:9189620; PIDN:AAA66994.1; PID:9189621
 A:Note: part of this sequence, including the amino end of the mature protein, was con
 R:Vinko, P.; Virkkunen, P.; Henttu, P.; Roiko, K.; Solin, T.; Hultala, M.L.
 FEBS Lett. 236, 275-281, 1988
 A:Title: Molecular cloning and sequence analysis of cDNA encoding human prostatic aci
 A:Reference number: S01331; MUID:86312881; PMID:2842184
 A:Accession: S01331
 A:Molecule type: mRNA
 A:Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-65, 'WMPTHPA', 74-211, 'A', 213-386 <VIH
 A:Cross-references: EMBL:X52174; NID:928321; PIDN:CAA36422.1; PID:928322
 A:Note: part of this sequence, including the amino end of the mature protein, was con
 R:Shartel, F.S.; Lee, H.; Leidenman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S
 Biochem. Biophys. Res. Commun. 160, 79-86, 1989
 A:Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein seq
 A:Reference number: A32419; MUID:89228054; PMID:2712834
 A:Accession: A32419
 A:Molecule type: mRNA
 A:Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-94, 'D', 96-115, 'R', 117-214, 'S', 216-293
 A:Cross-references: GB:M24902; NID:9189618; PIDN:AAA60022.1; PID:9189619
 A:Note: the authors translated the codons GAC for residue 95 as Glu, CGT for residue
 R:Tallor, P.G.; Govindan, M.V.; Patel, P.C.
 Nucleic Acids Res. 18, 4928, 1990
 A:Title: Nucleotide sequence of human prostatic acid phosphatase determined from a fu
 A:Reference number: S11147; MUID:90370491; PMID:2395659
 A:Accession: S11147
 A:Status: preliminary; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-14, 'AFAC', 20, 'CF', 23, 'C', 25-45, 'H', 47-65, 'RIMPTHPA', 74-138, 'E', 140-156
 A:Cross-references: EMBL:X53605; NID:935663; PIDN:CAA37673.1; PID:935684
 R:Banas, B.; Blaschke, D.; Filtter, F.; Hoetz, W.
 submitted to the EMBL Data Library, April 1993
 A:Description: Characterization of the promoter of the human prostatic acid phosphata
 A:Reference number: S38863
 A:Accession: S38863
 A:Molecule type: DNA
 A:Residues: 1-40 <BAN>
 A:Cross-references: EMBL:X71391; NID:9416530; PIDN:CAA50514.1; PID:9416531
 R:Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Blitt, E.; Porvari, K.; Taavitsalinen, P
 submitted to the EMBL Data Library, September 1993
 A:Description: Structural organization of human and rat prostate-specific acid phosph
 sequence in the human gene promoter.
 A:Reference number: S41251
 A:Accession: S41251
 A:Molecule type: DNA
 A:Residues: 1-40 <VIR>

A:Cross-references: EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:g439658
 R:Lee, H.; Chu, T.M.; Li, S.S.; Lee, C.
 Biochem. J. 277, 759-765, 1991
 A:Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.
 A:Reference number: S17042; MUID:91336999; PMID:1908222
 A:Accession: S17042
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 33-49 <LE>
 R:Benas, B.; Blaschke, D.; Fittler, F.; Hoertz, W.
 Biochim. Biophys. Acta 1217, 188-194, 1994
 A:Title: Analysis of the promoter of the human prostatic acid phosphatase gene.
 A:Reference number: S42730; MUID:94153995; PMID:8110833
 A:Accession: S42730
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-40 <BA>
 A:Cross-references: GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
 R:Sharief, F.S.; Li, S.S.
 Biochem. Mol. Biol. Int. 33, 561-565, 1994
 A:Title: Nucleotide sequence of human prostatic acid phosphatase ACP gene.
 A:Reference number: I37175; MUID:95038536; PMID:7951074
 A:Accession: I37175
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-386 <RES>
 A:Cross-references: EMBL:U07097; NID:g515995; PIDN:AB60640.1; PID:g515997
 C:Comment: This protein is synthesized under androgen regulation by epithelial cells of C:Genetics:
 A:Gene: GDB:ACPP
 A:Cross-references: GDB:119644; OMIM:171790
 A:Map position: 3q21.3-q25.2
 A:Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2
 C:Function:
 A:Description: catalyzes the hydrolysis of a wide range of phosphate esters
 C:Superfamily: mammalian acid phosphatase
 C:Keywords: glycoprotein; phosphotidine; phosphoprotein; phosphoric monoester hydrolysis
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:33-386/Product: acid phosphatase ACP #status experimental <MAT>
 F:44/Active site: Arg #status predicted
 F:94,220,333/Binding site: His (phosphotidine intermediate) #status predicted
 F:161-372,215-313,347-351/Disulfide bonds: #status experimental

Query Match 5.1%; Score 35; DB 1; Length 386;
 Best Local Similarity 100.0%; Pred. No. 3.4e-26;
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRAPLLARASLSIGFLFFLFWLDRSVLAKEL 35
 DB 1 MRAPLLARASLSIGFLFFLFWLDRSVLAKEL 35

RESULT 6
 A61632
 granulocyte-macrophage colony-stimulating factor precursor - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
 C:Accession: A61632
 R:O'Brien, P.M.; Rothel, J.S.; Seow, H.F.; Wood, P.R.
 Immunol. Cell Biol. 69, 51-55, 1991
 A:Title: Cloning and sequencing of the cDNA for ovine granulocyte-macrophage colony-stimulating factor
 A:Reference number: A61632; MUID:91331592; PMID:1869289
 A:Accession: A61632
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-144 <OAB>
 A:Cross-references: GB:X55991; NID:g6983759; PIDN:CAA39463.1; PID:g6983760
 C:Superfamily: granulocyte-macrophage colony-stimulating factor
 F:1-17/Domain: signal sequence #status predicted <SIG>

Query Match 3.2%; Score 22; DB 1; Length 144;
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 598 QEPICLTQRLRYKQGLRSLT 619
 DB 67 QEPICLTQRLRYKQGLRSLT 88

RESULT 7
 JH0469
 granulocyte-macrophage colony-stimulating factor precursor - sheep
 N:Alternate names: colony-stimulating factor 2; GM-CSF
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C:Accession: JH0469; S16730
 R:McInnes, C.J.; Haig, D.M.
 Gene 105, 275-279, 1991
 A:Title: Cloning and expression of a cDNA encoding ovine granulocyte-macrophage colony-stimulating factor
 A:Reference number: JH0469; MUID:92039044; PMID:1937025
 A:Accession: JH0469
 A:Molecule type: mRNA
 A:Residues: 1-144 <MC1>
 A:Cross-references: GB:X53561; NID:g1800; PIDN:CAA37632.1; PID:g1801
 C:Comment: This protein is a glycoprotein cytokine produced and secreted by various c
 C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-144/Product: granulocyte-macrophage colony-stimulating factor #status predicted
 F:44/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 3.2%; Score 22; DB 2; Length 144;
 Best Local Similarity 100.0%; Pred. No. 1.1e-13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 QEPICLTQRLRYKQGLRSLT 619
 DB 67 QEPICLTQRLRYKQGLRSLT 88

RESULT 8
 A44936
 granulocyte-macrophage colony-stimulating factor precursor - dog
 C:Species: Canis lupus familiaris (dog)
 C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999
 C:Accession: A44936
 R:Nash, R.A.; Schuening, F.; Appelbaum, F.; Hammond, W.P.; Boone, T.; Morris, C.F.; S
 Blood 78, 930-937, 1991
 A:Title: Molecular cloning and in vivo evaluation of canine granulocyte-macrophage co
 A:Reference number: A44936; MUID:91329842; PMID:1868252
 A:Accession: A44936
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-144 <NAS>
 A:Cross-references: GB:S49738; NID:g233566; PIDN:AB19466.1; PID:g233567
 A:Note: sequence extracted from NCBI backbone (NCBI:49738, NCBI:P:49739)
 C:Superfamily: granulocyte-macrophage colony-stimulating factor
 F:1-17/Domain: signal sequence #status predicted <SIG>

Query Match 1.6%; Score 11; DB 2; Length 144;
 Best Local Similarity 100.0%; Pred. No. 0.0095;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 631 HYKHCPTPE 641
 DB 100 HYKHCPTPE 110

RESULT 9
 F080GM
 granulocyte-macrophage colony-stimulating factor precursor - bovine
 N:Alternate names: colony-stimulating factor 2; GM-CSF
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 27-Jan-1995
 C:Accession: J10037

R.Malliszewski, C.R.; Schoenborn, M.A.; Cerretti, D.P.; Wignall, J.M.; Picha, K.S.; Cosma
Mol. Immunol. 25, 843-850, 1988
A:Title: Bovine GM-CSF: molecular cloning and biological activity of the recombinant pro
A:Reference number: J10037; MUID:89096971; PMID:3062386
A:Accession: J10037

A:Molecule type: mRNA
A:Residues: 1-143 <MAL>
C:Comment: This glycoprotein induces granulocyte, macrophage, and eosinophil colony form
C:Superfamily: granulocyte-macrophage colony-stimulating factor
C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-143/Product: granulocyte-macrophage colony-stimulating factor #status predicted <MA
F:44,54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 1.4% Score 10: DB 1: Length 143;
Best Local Similarity 100.0%; Pred. No. 0.092;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 598 QEPCTCQTRL 607
DB 66 QEPCTCQTRL 75

RESULT 10
A42032
epidermal growth factor receptor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
C:Accession: A42032
R.Flickinger, T.W.; Maibach, N.J.; Kung, H.J.
Mol. Cell. Biol. 12, 883-893, 1992
A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t
A:Reference number: A42032; MUID:92123214; PMID:1732751
A:Accession: A42032

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-527 <FLI>
A:Cross-references: GB:M77637; NID:9211737; PIDN:AAA48759.1; PID:9211738
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone (NCBIN:76692, NCBI:76693)
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; growth factor receptor

Query Match 1.3% Score 9: DB 2: Length 527;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266
DB 245 QCAAGCTGP 253

RESULT 11
A36325
epidermal growth factor receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 10-Oct-1997
C:Accession: A36325
R.Petich, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; EARP, H.S.
Mol. Cell. Biol. 10, 2973-2982, 1990
A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b
A:Reference number: A36325; MUID:90258888; PMID:2342466
A:Accession: A36325
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <PET>
A:Cross-references: GB:M37394
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 1.3% Score 9: DB 2: Length 644;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266
DB 235 QCAAGCTGP 243

RESULT 12
G0HUE
epidermal growth factor receptor precursor - human
N:Contents: protein-tyrosine kinase (EC 2.7.1.112) erbb
C:Species: Homo sapiens (man)
C:Date: 15-Nov-1984 #sequence_revision 27-Nov-1985 #text_change 11-Jun-1999
C:Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05283; A60143;
R.Ullrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.
ig, P.H.

Nature 309, 418-425, 1984
A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression
A:Reference number: A00641; MUID:84219729; PMID:6328312
A:Accession: A00641

A:Molecule type: mRNA
A:Residues: 1-1210 <ULL>
A:Cross-references: EMBL:X00588; NID:931113; PIDN:CAA25240.1; PID:9757924
A:Note: the authors translated the codon AAG for residue 340 as Asn
R.Ishii, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
A:Title: Characterization and sequence of the promoter region of the human epidermal
A:Reference number: A25772; MUID:85270438; PMID:2991899
A:Accession: A25772

A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-29 <KSH>
A:Cross-references: GB:M11234; NID:9181981; PIDN:AAA52370.1; PID:9553272
R.Haley, J.; White, N.; Bennett, P.; Kinchington, D.; Ullrich, A.; Waterfield, M.
Oncogene Res. 1, 375-396, 1987
A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification
A:Reference number: S30024; MUID:88217333; PMID:3329716
A:Accession: S30024

A:Molecule type: DNA
A:Residues: 1-29 <HA2>
A:Cross-references: EMBL:X06370; NID:931118; PIDN:CAA29668.1; PID:931119
R.Haley, J.D.; Waterfield, M.D.
J. Biol. Chem. 266, 1746-1753, 1991
A:Title: Contributory effects of de Novo transcription and premature transcript termi
A:Reference number: A38672; MUID:91107677; PMID:1988448
A:Accession: A38672

A:Molecule type: DNA
A:Residues: 1-29 <HAL>
A:Cross-references: GB:M38425; NID:9181977; PIDN:AAA63171.1; PID:9553271
A:Experimental source: carcinoma cell line A431-7
R.Xu, Y.; Ishii, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;
Nature 309, 806-810, 1984
A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN
A:Reference number: A00642; MUID:84245835; PMID:6330563
A:Accession: A00642

A:Molecule type: mRNA
A:Residues: 'RCAMRRA', 150-187, 'KSVIQAV', 195, 'M', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-
'798-799, 'ND', 802-811, 'R', 813-942 <XU>
A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF re
R.Ullrich, C.R.; Chen, W.S.; Kruliger, W.; Stokarsky, L.S.; Weber, W.; Evans, R.M.; Verma,
Science 224, 843-848, 1984
A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplificati
A:Reference number: A43615; MUID:84196372; PMID:6326261
A:Accession: A43615

A:Molecule type: mRNA
A:Residues: 713-964 <LIN>
A:Experimental source: epidermoid carcinoma cell line A431
R.Stimmen, F.A.; Gope, M.L.; Schultz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
Biochem. Biophys. Res. Commun. 124, 125-132, 1984
A:Reference number: A23062; MUID:85046483; PMID:6093780
A:Accession: A23062

A:Molecule type: mRNA
A:Residues: 1028-1210 <SIM>
R.Weber, W.; Gill, G.N.; Speiss, J.

Science 224, 294-297, 1984
 A:Reference number: A05281; MUID:84172183; PMID:6324343
 A:Accession: A05281
 A:Molecule type: protein
 R:Residues: 25-30,'S',32-51,45A-46T <MEB>
 J. Biol. Chem. 260, 5205-5208, 1985
 A:Title: Identification of residues in the nucleotide binding site of the epidermal growth factor receptor
 A:Reference number: A60143; MUID:85182650; PMID:2985580
 A:Accession: A60143
 A:Molecule type: protein
 A:Residues: 740-744,'X',746-747 <RUS>
 R:Mroczkowski, B.; Mosig, G.; Cohen, S.
 Nature 309, 270-273, 1984
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and superoxide dismutase
 A:Reference number: A38023; MUID:84191554; PMID:6325948
 A:Accession: A38023
 A:Contents: annotation; receptor activity
 A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
 R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Weisn, J.B.; Chang, C.P.; Walton, G.M.; Der, C.J.
 Cell 59, 33-43, 1989
 A:Title: Functional independence of the epidermal growth factor receptor from a domain X
 A:Reference number: A33331; MUID:90003233; PMID:2750960
 A:Accession: A33331
 A:Contents: annotation; internalization signal
 C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-receptor
 C:Genetics:
 A:Gene: GDB:EGFR
 A:Cross-references: GDB:120610; OMIM:131550
 A:Map position: 7p12.3-7p12.1
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:23-1210/Product: EGF receptor #status predicted <MAT>
 F:23-665/Domain: extracellular #status predicted <EXT>
 F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>
 F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>
 F:666-668/Domain: transmembrane #status predicted <TM>
 F:669-1210/Domain: intracellular #status predicted <INT>
 F:710-973/Domain: protein kinase homology <KIN>
 F:718-726/Region: protein kinase ATP-binding motif
 F:999-1046/Region: coated-pit mediated internalization signal
 F:1047-1210/Region: inhibitory
 F:128,175,352,413,444,528,603/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:745/Active site: Lys #status experimental
 Query Match 1.38; Score 9; DB 1; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266
 DB 235 QCAAGCTGP 243

RESULT 13
 A53183
 epidermal growth factor receptor precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
 C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
 R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A.;
 Genes Dev. 8, 399-413, 1994
 A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF receptor
 A:Reference number: A53183; MUID:94170986; PMID:8125255
 A:Accession: A53183
 A:Molecule type: mRNA
 A:Residues: 1-1210 <UE>
 A:Cross-references: GB:U03425
 R:Aviv, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
 Oncogene 6, 673-676, 1991
 A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding site
 A:Reference number: A43818; MUID:91232866; PMID:2030916
 A:Accession: A43818
 A:Molecule type: mRNA

A:Residues: 1-714 <AVI>
 A:Cross-references: GB:X59698
 R:Elisinger, D.P.; Serrero, G.
 submitted to the EMBL Data Library, June 1992
 A:Reference number: S24942
 A:Accession: S24942
 A:Molecule type: mRNA
 A:Residues: 969-971,'K',973-1115,'D' <EIS>
 A:Cross-references: EMBL:Z12608
 R:Heisermann, G.J.; Gill, G.N.
 J. Biol. Chem. 263, 13152-13158, 1988
 A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylate
 A:Reference number: A28941; MUID:88330814; PMID:3138233
 A:Accession: A28941
 A:Molecule type: protein
 A:Residues: 669-674,'X',696-704,'V',706-707,989-992,'XX',995-996,'X',998-1000,1002-10
 R:Hbbs, M.L.; Dunn, A.R.; Alexander, W.S.
 submitted to the EMBL Data Library, April 1994
 A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Recept
 A:Reference number: S45325
 A:Accession: S45325
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-971,'K',973-1210 <VER>
 A:Cross-references: EMBL:X78987; NID:9488830; PIDN:CAA55587.1; PID:9488831
 R:Paria, B.C.; Das, S.K.; Andrews, G.K.; Dey, S.K.
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
 A:Title: Expression of the epidermal growth factor receptor gene is regulated in mos
 A:Reference number: I49643; MUID:93126380; PMID:7678348
 A:Accession: I49643
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 12-20,22-132 <RES>
 A:Cross-references: GB:I06864; NID:9193001; PIDN:AAA53029.1; PID:9567201
 C:Genetics:
 A:Gene: EGFR
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phospho
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:648-670/Domain: transmembrane #status predicted <TM>
 F:712-977/Domain: protein kinase homology <KIN>
 F:720-728/Region: protein kinase ATP-binding motif
 F:680,695/Binding site: phosphate (Thr) (covalent) #status experimental
 F:697,1070,1071/Binding site: phosphate (Ser) (covalent) #status experimental
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental
 Query Match 1.38; Score 9; DB 2; Length 1210;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266
 DB 235 QCAAGCTGP 243

RESULT 14
 T4CHLV
 epidermal growth factor receptor precursor - chicken
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
 C:Species: Gallus gallus (chicken)
 C:Date: 28-Feb-1986 #sequence_revision 05-May-1995 #text_change 04-Feb-2000
 C:Accession: A27720; A00643
 R:Lax, I.; Johnson, A.; Hawk, R.; Sap, J.; Bellot, F.; Winkler, M.; Ullrich, A.; Venn
 Mol. Cell. Biol. 8, 1970-1978, 1988
 A:Title: Chicken epidermal growth factor (EGF) receptor: cDNA cloning, expression in
 A:Reference number: A27720; MUID:88261272; PMID:3260329
 A:Accession: A27720
 A:Molecule type: mRNA
 A:Residues: 1-1223 <LAX>
 A:Cross-references: GB:M20386
 R:Nilsen, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Raines

Cell 41, 719-726, 1985

A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and pro

A:Reference number: A00643; MUID:85228222; PMID:2988784

A:Accession: A00643

A:Molecule type: mRNA

A:Residues: 585-1223 <N1L>

A:Cross-references: GB:M10066

C:Genetics:

A:Gene: erbB

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor

pecific protein kinase

F:1-30/Domain: signal sequence #status predicted <SIG>

F:31-1223/Product: epidermal growth factor receptor #status predicted <MAT>

F:31-654/Domain: extracellular #status predicted <EXT>

F:81-307/Domain: EGF receptor extracellular domain repeat <EB1>

F:397-610/Domain: EGF receptor extracellular domain repeat <EB2>

F:655-677/Domain: transmembrane #status predicted <TM>

F:719-984/Domain: intracellular #status predicted <INT>

F:727-735/Region: protein kinase ATP-binding motif

F:136-202,280,361,370,422,575,580,615,635/Binding site: carbohydrate (Thr) (covalent) #

F:192,650/Binding site: carbohydrate (Ser) (covalent) #status predicted

F:687/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F:754/Active site: Lys #status predicted

F:1100,1183,1208/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Query Match

Best Local Similarity 1.3%; Score 9; DB 1; Length 1223;

Best Local Similarity 100.0%; Pred. No. 6.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 258 QCAAGCTGP 266

DB 242 QCAAGCTGP 250

RESULT 15

I46269

granulocyte-macrophage colony stimulating factor - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 14-Feb-1997 #sequence-revision 14-Feb-1997 #text-change 16-Jul-1999

C:Accession: I46269

R:Smith, L.R.; Lundeen, K.A.; Diveley, J.P.; Carlo, D.J.; Brostoff, S.W.

Immunogenetics 39, 80, 1994

A:Title: Nucleotide sequence of the Lewis rat granulocyte-macrophage colony stimulating

A:Reference number: I46269; MUID:94041474; PMID:8225444

A:Accession: I46269

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-127 <SM1>

A:Cross-references: EMBL:U00620; NID:g392779; PIDN:AA18281.1; PID:g392780

C:Superfamily: granulocyte-macrophage colony-stimulating factor

Query Match

Best Local Similarity 1.2%; Score 8; DB 2; Length 127;

Best Local Similarity 100.0%; Pred. No. 8;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 609 LYKQGLRG 616

DB 61 LYKQGLRG 68

Search completed: April 28, 2003, 13:46:15
Job time : 26 secs

THIS PAGE BLANK (USPTO)

DR WPI: 2001-662965/76.
 DR N-PSDB: AAD21565.
 PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 is useful for the treatment of associated cancer associated -
 PS Claim 7; Page 26; 59pp: English.
 XX The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC hbm-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC intracellular domain, an Ala Ala linker, a mature human granulocyte-
 CC macrophage colony stimulating factor (GM-CSF) sequence and a
 CC C-terminal tag.
 SQ Sequence 690 AA:
 Query Match 100.0%; Score 690; DB 22; Length 690;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOVCSTDMKRLRPASPER 60
 DB 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOVCSTDMKRLRPASPER 60
 QY 61 HLDMLRLHYOGCOVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROVFLORLRI 120
 DB 61 HLDMLRLHYOGCOVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROVFLORLRI 120
 QY 121 VRGTQLEFEDNVYALVLDNGDPLNNTPTVTGASPGGLRELRLSLTEILKGGVLLIQRRPOL 180
 DB 121 VRGTQLEFEDNVYALVLDNGDPLNNTPTVTGASPGGLRELRLSLTEILKGGVLLIQRRPOL 180
 QY 181 CYODTILMKDIFHKNNDLALTLIDTNSRACHSPCKSGRCGSESDCQSTIRTVCA 240
 DB 181 CYODTILMKDIFHKNNDLALTLIDTNSRACHSPCKSGRCGSESDCQSTIRTVCA 240
 QY 241 GGCARCKGRLPTDCHEQCAAGCTGPRKSDCLAFHNHSGICELHCPALVTYNTDFEES 300
 DB 241 GGCARCKGRLPTDCHEQCAAGCTGPRKSDCLAFHNHSGICELHCPALVTYNTDFEES 300
 QY 301 MPNDEGRYTFGASCVTACPYNYLSTDVSGAGCMVHHHRSSSTRSGGDLTLGLPSEE 360
 DB 301 MPNDEGRYTFGASCVTACPYNYLSTDVSGAGCMVHHHRSSSTRSGGDLTLGLPSEE 360
 QY 361 EAPRSPLAPBEGAGSDVFDGDLGMAKGLQSLPTHRPSPLQRISEPTPLPLPSETGYV 420
 DB 361 EAPRSPLAPBEGAGSDVFDGDLGMAKGLQSLPTHRPSPLQRISEPTPLPLPSETGYV 420
 QY 421 APLTCSPOPEYVNPQVRPQPPSPREGRLPAARPAAGATLEAKTLGSKGNVYDVAFG 480
 DB 421 APLTCSPOPEYVNPQVRPQPPSPREGRLPAARPAAGATLEAKTLGSKGNVYDVAFG 480
 QY 481 GAVENPEYLTTPQGAAPQHPPEAFSPAFDNLVYWDODPEPGAPSTFGKPTAENPEY 540
 DB 481 GAVENPEYLTTPQGAAPQHPPEAFSPAFDNLVYWDODPEPGAPSTFGKPTAENPEY 540
 QY 541 LGIDVPAAARARSPSTQGWENHNAIOEARRLNLSRDAEKNNEVEYISEMFDQEP 600
 DB 541 LGIDVPAAARARSPSTQGWENHNAIOEARRLNLSRDAEKNNEVEYISEMFDQEP 600
 QY 601 TCIQTRLELTKOGLRSLTKLGPITMASHYKOHCPPTETSCATQIITFESKKNLKD 660
 DB 601 TCIQTRLELTKOGLRSLTKLGPITMASHYKOHCPPTETSCATQIITFESKKNLKD 660

QY 661 FLVYIPDCMEYOEAGARRPAAANNHNN 690
 DB 661 FLVYIPDCMEYOEAGARRPAAANNHNN 690
 RESULT 2
 ID AAE13108 standard; Protein; 555 AA.
 AC AAE13108;
 DT 28-JAN-2002 (first entry)
 DE Human HER500 fusion protein construct.
 XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW Immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human;
 KW HER-2 protein; HER500 fusion protein.
 XX Chimeric - Homo sapiens.
 OS Chimeric - Synthetic.
 PN WO200174855-A2.
 PD 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US10515.
 XX 30-MAR-2000; 2000US-193504P.
 PR (DENB-) DENDREON CORP.
 PA Laus R, Vidovic D, Graddis T;
 DR WPI: 2001-662965/76.
 DR N-PSDB: AAD21564.
 PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 is useful for the treatment of associated cancer associated -
 PS Claim 7; Page 26; 59pp: English.
 XX The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC fusion protein construct which comprises human PAP signal
 CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
 CC sequence, mature HER-2 membrane distal extracellular and intracellular
 CC domains and a C-terminal tag.
 SQ Sequence 555 AA:
 Query Match 79.6%; Score 549; DB 22; Length 555;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOVCSTDMKRLRPASPER 60
 DB 1 MRAAPLLARAASLSLGFLLFFWLDRLSVLAKLARGAASSTOVCSTDMKRLRPASPER 60
 QY 61 HLDMLRLHYOGCOVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROVFLORLRI 120
 DB 61 HLDMLRLHYOGCOVYVGNLELTYLPTNASLSFLDIOEVGYVLIANOVROVFLORLRI 120

PD 11-OCT-2001.
 XX 30-MAR-2001; 2001MO-US10515.
 PF 30-MAR-2000; 2000US-193504P.
 PR (DEND-) DENDREON CORP.
 XX
 PA Laus R, Vidovic D, Graddis T;
 PI WPI; 2001-662965/76.
 DR N-PSDB; AAD21566.
 XX
 PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 is useful for the treatment of associated cancer associated -
 XX
 PS Claim 7; Page 26; 59pp: English.
 CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC fusion protein construct which comprises human PAP signal
 CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
 CC sequence, mature HER-2 membrane distal extracellular domain, an
 CC Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
 CC HER-2 membrane distal intracellular domain and a C-terminal tag.
 XX
 SQ Sequence 564 AA:
 Query Match 47.7%; Score 329; DB 22; Length 564;
 Best Local Similarity 100.0%; Pred. No. 3.9e-293;
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAAPLLARAAASISLGFLLFFWMDRSVLAKEARGAASSTOVCSTGDMKRLRPASPET 60
 DB 1 MRAAPLLARAAASISLGFLLFFWMDRSVLAKEARGAASSTOVCSTGDMKRLRPASPET 60
 QY 61 HLDMLRLHYGCGVYVGNLEETVLPNTASISFLDDIOEVGYVLIANOVROYVQLRLRI 120
 DB 61 HLDMLRLHYGCGVYVGNLEETVLPNTASISFLDDIOEVGYVLIANOVROYVQLRLRI 120
 QY 121 VRGTFEDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180
 DB 121 VRGTFEDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180
 QY 181 CYODTILMKDIFHKNNGLATLTLIDTNSRACHPCSPCKGSRGWESSEDCQSLTRTVCA 240
 DB 181 CYODTILMKDIFHKNNGLATLTLIDTNSRACHPCSPCKGSRGWESSEDCQSLTRTVCA 240
 QY 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLACLPNHSGICELHPALVYNTDTFFES 300
 DB 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLACLPNHSGICELHPALVYNTDTFFES 300
 QY 301 MPNPEGRTYFGASCVTACPYNYLSTDVGS 329
 DB 301 MPNPEGRTYFGASCVTACPYNYLSTDVGS 329

RESULT 5
 AAE13111
 ID AAE13111 standard; Protein; 697 AA.
 XX
 AC AAE13111;
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human HER500-rgm-CSF fusion construct comprising OVA-derived peptide.

XX
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
 KW ovalbumin-derived octapeptide; OVA; rat; HER500-rgm-CSF fusion protein.
 XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Rattus norvegicus.
 OS Chimeric - Unidentified.
 XX
 PN WO200174855-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US10515.
 XX
 PR 30-MAR-2000; 2000US-193504P.
 XX
 PA (DEND-) DENDREON CORP.
 PI Laus R, Vidovic D, Graddis T;
 DR WPI; 2001-662965/76.
 DR N-PSDB; AAD21566.
 XX
 PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 is useful for the treatment of associated cancer associated -
 XX
 PS Claim 7; Page 27; 59pp: English.
 CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC rgm-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC signal sequence, mature HER-2 membrane distal extracellular domain,
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
 CC HER-2 membrane distal intracellular domain, an Ala Ala linker, a mature
 CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
 CC and a C-terminal tag.
 XX
 SQ Sequence 697 AA:
 Query Match 47.7%; Score 329; DB 22; Length 697;
 Best Local Similarity 100.0%; Pred. No. 4.6e-293;
 Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRAAPLLARAAASISLGFLLFFWMDRSVLAKEARGAASSTOVCSTGDMKRLRPASPET 60
 DB 1 MRAAPLLARAAASISLGFLLFFWMDRSVLAKEARGAASSTOVCSTGDMKRLRPASPET 60
 QY 61 HLDMLRLHYGCGVYVGNLEETVLPNTASISFLDDIOEVGYVLIANOVROYVQLRLRI 120
 DB 61 HLDMLRLHYGCGVYVGNLEETVLPNTASISFLDDIOEVGYVLIANOVROYVQLRLRI 120
 QY 121 VRGTFEDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180
 DB 121 VRGTFEDNRYALAVLDNGDPLNNTTPVTGASPGGLRELDRLSTELIKGCVLIQRNPOL 180
 QY 181 CYODTILMKDIFHKNNGLATLTLIDTNSRACHPCSPCKGSRGWESSEDCQSLTRTVCA 240
 DB 181 CYODTILMKDIFHKNNGLATLTLIDTNSRACHPCSPCKGSRGWESSEDCQSLTRTVCA 240
 QY 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLACLPNHSGICELHPALVYNTDTFFES 300

Db 241 GGCACCKCPPLPDCCHCCAGCTGPKHSDCLAHFHNHSGICEHCPALVTYNTDTEES 300
OY 301 MPNPEGRTFGASCVTACPYNYLSTDVGS 329
Db 301 MPNPEGRTFGASCVTACPYNYLSTDVGS 329

RESULT 6

AAB60408 standard; Protein: 645 AA.

XX AAB60408;

DT 24-APR-2001 (first entry)

XX Human ErbB2 oncoprotein, SEQ ID NO:13.

XX Anti-ErbB2 monoclonal antibody 2C4; HER2; mouse; murine; humanised; VL;
XX light chain variable region; cancer; cytosolic; EGFR-expressing cancer;
XX epidermal growth factor receptor; colon cancer; rectal cancer; tumour;
XX colorectal cancer; non-small cell lung cancer; metastatic breast cancer;
XX affinity purification.

XX Homo sapiens.

XX WO200100245-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-US17366.

XX 25-JUN-1999; 99US-0141316.

XX (GETH) GENENTECH INC.

XX Adams CW, Presta LG, Sliwkowsky M;

XX WPI; 2001-080862/09.

XX Treating cancer in a human, where the cancer expresses epidermal growth
XX factor receptor (EGFR), comprises administering an antibody which binds
XX ErbB2 -

XX Example 1; Fig 1A; 89pp; English.

XX The invention relates to a method for treating cancer in a human patient,
XX wherein the cancer expresses epidermal growth factor receptor (EGFR),
XX comprising administering an antibody which binds ErbB2 (HER2; AAB60408).
XX In particular, the anti-ErbB2 antibody is the murine monoclonal antibody
XX 2C4 (AAB60396, AAB60397) or a humanised version of 2C4 (AAB60398,
XX AAB60399). The invention also encompasses an isolated nucleic acid
XX encoding a humanised ErbB2-binding antibody; vectors and host cells
XX comprising such nucleic acids; the recombinant production of a humanised
XX ErbB2-binding antibody; and an immunocjugate comprising a humanised
XX ErbB2-binding antibody and a cytotoxic drug. The ErbB2-binding antibodies
XX act by antagonising ErbB receptors, and as inhibitors of transforming
XX growth factor alpha (TGF-alpha)-activated mitogen activated protein
XX kinase (MAPK). The method of the invention is used for treating cancer,
XX especially colon cancer, rectal cancer, colorectal cancer, lung cancer
XX (especially non-small cell lung cancer), or breast cancer (especially
XX metastatic breast cancer). The antibodies may also have non-therapeutic
XX uses e.g., as affinity purification agents. Using an antibody which binds
XX to ErbB2 to treat cancer is preferable to the use of EGFR-targeted
XX drugs, as EGFR is also highly expressed in other tissues such as the
XX liver and skin, where the active drug will also bind, with skin toxicity
XX having been observed for EGFR-targeted drugs. Antibodies which bind
XX ErbB2 are anticipated to have a better safety profile than such drugs.
XX The present sequence represents human ErbB2.

XX Sequence 645 AA;

Query Match 42.3%, Score 292; DB 22; Length 645;

Best Local Similarity 100.0%; Pred. No. 4e-259;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAASQVCTGDMKRLPASPETHIDMLRHLYCGGVVGNLELTYLPTNASTSLDIOIQ 97
Db 19 GAASQVCTGDMKRLPASPETHIDMLRHLYCGGVVGNLELTYLPTNASTSLDIOIQ 78
OY 98 EVQGYVLLAHNOVROVPLQRLRYVGTOLFEDNYALAVLDNGDPLNNTPTVTSAPGSLR 157
Db 79 EVQGYVLLAHNOVROVPLQRLRYVGTOLFEDNYALAVLDNGDPLNNTPTVTSAPGSLR 138
OY 158 ELQRLSLTEILKGVLIQNPOLCYQDTTLWKDIFHKNNOALATLIDNRSRACHPCSPM 217
Db 139 ELQRLSLTEILKGVLIQNPOLCYQDTTLWKDIFHKNNOALATLIDNRSRACHPCSPM 198
OY 218 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGRLPDCCHCCAGCTGPKHSDCLAHF 277
Db 199 CKGSRCKWESSSDCOSLRTVTCAGGCARCKGRLPDCCHCCAGCTGPKHSDCLAHF 258
OY 278 NMSGICEHCPALVTYNTDTEESMPNPEGRTFGASCVTACPYNYLSTDVGS 329
Db 259 NMSGICEHCPALVTYNTDTEESMPNPEGRTFGASCVTACPYNYLSTDVGS 310

RESULT 7

AAB61593 standard; protein: 645 AA.

XX AAB61593;

DT 04-APR-2001 (first entry)

XX Human ErbB2 extracellular domain.

XX Human; ErbB2; cytosolic; prostate cancer; receptor tyrosine kinase;
XX antibody; ErbB receptor; monoclonal antibody 2C4; variable light chain.

XX Homo sapiens.

XX WO200100238-A1.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-US17423.

XX 25-JUN-1999; 99US-0141315.

XX (GETH) GENENTECH INC.

XX (SLOK) SLOAN KETTERING INST CANCER RES.

XX Agus DB, Scher HI, Sliwkowski M;

XX WPI; 2001-159131/16.

XX Treating prostate cancer in a human comprises administering an antibody
XX which binds ErbB2 and blocks ligand activation of an ErbB receptor -
XX Disclosure; Fig 1; 93pp; English.

XX The ErbB family of receptor tyrosine kinases are important mediators of
XX cell growth, differentiation and survival. The receptor family includes
XX four distinct members including Epidermal Growth Factor Receptor (EGFR or
XX ErbB1), HER2 (ErbB2 or p185^{neu}), HER3 (ErbB3) and Her4 (ErbB4 or tyro2).
XX The present invention relates to a method for treating prostate cancer.
XX The method comprises administering an antibody which binds ErbB2 and
XX blocks ligand activation of an ErbB receptor. Preferably, the antibody
XX blocks binding of monoclonal antibody 2C4 to ErbB2 and/or blocks
XX TGF-alpha activation of mitogen-activated protein kinase (MAPK). The
XX present sequence is the extracellular domain of human ErbB2.

XX Sequence 645 AA;

Query Match 42.3%, Score 292; DB 22; Length 645;

CC Intracellular domain that also shows homology to EGFR. Its
CC overexpression correlates with a poor prognosis in breast and
CC ovarian cancers. The invention provides Her-2/neu fusion
CC proteins, nucleic acids encoding them, viral vectors, and vaccines
CC comprising the fusion proteins or nucleic acid molecules. In
CC preferred fusion proteins, the extracellular domain of a Her-2/neu
CC protein is fused to a Her-2/neu intracellular domain or
CC phosphorylation domain (or its delta fragment). An immune
CC response to Her-2/neu protein is elicited or enhanced by
CC administering the fusion protein in the form of a vaccine, or by
CC transfecting cells of an animal *ex vivo* with a nucleic acid
CC encoding the fusion protein, and delivering the transfected cells
CC to the animal. The fusion proteins, nucleic acids, and isolated
CC specific T-cells are useful for inhibiting the development of a
CC cancer, especially breast, ovarian, colon, lung or prostate cancer
CC in a patient. T cells that specifically react with a Her-2/neu
CC fusion protein can be used to remove tumour cells from a sample in
CC order to inhibit the development of cancer in a patient.

XX Sequence 653 AA;

Query Match 42.3%; Score 292; DB 23; Length 653;

Best Local Similarity 100.0%; Pred. No. 4.1e-259; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAATGCTGDMKRLPASPETHLMDLRHLHYGCGVVGMLDLYLPTNLSLFLDIO 97
DB 19 GAATGCTGDMKRLPASPETHLMDLRHLHYGCGVVGMLDLYLPTNLSLFLDIO 78
OY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDGDDPLNTPPTGASPGGLR 157
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDGDDPLNTPPTGASPGGLR 138
OY 158 ELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198
OY 218 CKGRKMGESSEDCOSLTRTYCAGGACARCKGRLPTDCHECCAGACTGPKHSDCLACHF 277
DB 199 CKGRKMGESSEDCOSLTRTYCAGGACARCKGRLPTDCHECCAGACTGPKHSDCLACHF 258
OY 278 NHSGICELHCPALVYNTDTEFSMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
DB 259 NHSGICELHCPALVYNTDTEFSMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 10

AAB21204

ID AAB21204 standard; protein; 712 AA.

XX AAB21204;

XX 12-JAN-2001 (first entry)

XX Human HER-2/neu fusion protein.

XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;

XX breast cancer; prostate cancer; ovarian cancer; lung cancer;

XX colon cancer; fusion protein.

XX Homo sapiens.

XX Synthetic.

XX WO200044899-A1.

XX 03-AUG-2000.

XX 28-JAN-2000; 2000WO-US02164.

XX 29-JAN-1999; 99US-0117976.

XX (CORI-) CORIAX CORP.

XX (SMIK) SMITHKLINE BEECHAM.

XX Cheever MA, Gheysen D;
XX WPI: 2000-505976/45.
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
XX useful for vaccinating against breast, ovarian, colon, lung and
XX prostate cancers -
XX Claim 27; Fig 13; 128pp; English.

XX The present sequence is a fusion protein comprising the extracellular
XX domain and a preferred portion of the phosphorylation domain of the human
XX HER-2/neu protein. HER-2/neu is a member of the tyrosine kinase family of
XX receptor-like glycoproteins and shows homology to the epidermal growth
XX factor receptor (EGFR). It probably plays a part in cell growth and/or
XX differentiation. The HER-2/neu gene is an oncogene. HER-2/neu fusion
XX proteins may be used to treat or prevent cancer by eliciting or enhancing
XX an immune response to the HER-2/neu protein. They may be used to treat
XX malignancies such as breast, ovarian, colon, lung and prostate cancers,
XX and may be used as an antigen to vaccinate against these neoplasias.

XX Sequence 712 AA;

Query Match 42.3%; Score 292; DB 21; Length 712;

Best Local Similarity 100.0%; Pred. No. 4.4e-259; Mismatches 0; Indels 0; Gaps 0;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 GAATGCTGDMKRLPASPETHLMDLRHLHYGCGVVGMLDLYLPTNLSLFLDIO 97
DB 19 GAATGCTGDMKRLPASPETHLMDLRHLHYGCGVVGMLDLYLPTNLSLFLDIO 78
OY 98 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDGDDPLNTPPTGASPGGLR 157
DB 79 EVGGVLLAHNOVROVPLQRLRIYRGTLFEDNYALAVLDGDDPLNTPPTGASPGGLR 138
OY 158 ELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGVLLIQRNPOLCYODTILMKDIFHKNNQALATLIDNRSRACHPCSPM 198
OY 218 CKGRKMGESSEDCOSLTRTYCAGGACARCKGRLPTDCHECCAGACTGPKHSDCLACHF 277
DB 199 CKGRKMGESSEDCOSLTRTYCAGGACARCKGRLPTDCHECCAGACTGPKHSDCLACHF 258
OY 278 NHSGICELHCPALVYNTDTEFSMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
DB 259 NHSGICELHCPALVYNTDTEFSMPNPEGRTYFGASCVTACPYNYLSTDVGS 310

RESULT 11

AAM51149

ID AAM51149 standard; protein; 712 AA.

XX AAM51149;

XX 17-JUN-2002 (first entry)

XX Her-2/neu extracellular domain-delta-phosphorylation domain fusion.

XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;

XX tyrosine kinase; receptor; c-erbB2; gene therapy.

XX Homo sapiens.

XX Key

XX Domain

XX Domain

XX Domain

XX Domain

XX Domain

XX Domain

Location/Qualifiers

1..653

/note="extracellular domain"

654..712

/note="phosphorylation domain fragment"

WO200212341-A2.

14-FEB-2002.

XX 03-AUG-2001; 2001MO-US24283.
 PF 03-AUG-2000; 2000US-0632507.
 XX
 PR
 XX
 PA (CORI-) CORIXA CORP.
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 PI
 PI Cheever MA, Gheysen D;
 XX
 XX WPI; 2002-241743/29.
 DR
 XX
 PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain
 XX
 XX
 PS Claim 37; Fig 13; 141pp; English.
 CC The present sequence is that of a fusion protein between the
 CC extracellular domain and a fragment (DeltapD) of the phosphorylation
 CC domain of human Her-2/neu (see AAM51143), an oncogenic self-protein
 CC and target for anti-cancer vaccines. The fusion protein can be
 CC obtained by recombinant DNA methods. Her-2/neu overexpression
 CC correlates with a poor prognosis in breast and ovarian cancers.
 CC The invention provides Her-2/neu fusion proteins, nucleic acids
 CC encoding them, viral vectors, and vaccines comprising the fusion
 CC proteins or nucleic acid molecules. In preferred fusion proteins,
 CC the extracellular domain of Her-2/neu is fused to a Her-2/neu
 CC intracellular domain or phosphorylation domain (or its DeltapD
 CC fragment). An immune response to Her-2/neu protein is elicited or
 CC enhanced by administering the fusion protein in the form of a vaccine,
 CC or by transfecting cells of an animal ex vivo with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.
 CC
 CC
 SO Sequence 712 AA:
 Query Match 42.3%; Score 292; DB 23; Length 712;
 Best Local Similarity 100.0%; Pred. No. 4.4e-259;
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 38 GAASTOYCTGDMKRLRPASPEHLDMLRHLYGCGVVGNGLEETVPTNASLSFLDIO 97
 DB 19 GAASTOYCTGDMKRLRPASPEHLDMLRHLYGCGVVGNGLEETVPTNASLSFLDIO 78
 QY 98 EVGGYVLIANNOVROVPLQRLIRVGTQLFEDNVALAVLNDGDP LNNTPTVGTASPGGLR 157
 DB 79 EVGGYVLIANNOVROVPLQRLIRVGTQLFEDNVALAVLNDGDP LNNTPTVGTASPGGLR 138
 QY 158 ELQIRSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPM 217
 DB 139 ELQIRSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPM 198
 QY 218 CKGSRGCGESSEDCOSLTRIVCAGGCARCKGRLPTDCCHQCAAGCGPRHSDCLACLHF 277
 DB 199 CKGSRGCGESSEDCOSLTRIVCAGGCARCKGRLPTDCCHQCAAGCGPRHSDCLACLHF 258
 QY 278 NMSGICELHCPALVTYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 329
 DB 259 NMSGICELHCPALVTYNTDTFESMPNPEGRTYFGASCVTACPYNYLSTDVGS 310
 RESULT 12
 AAM19764
 ID AAM19764 standard; Protein; 782 AA.
 XX
 AC AAM19764;

XX 17-SEP-1997 (first entry)
 DT
 XX
 DE Her2-GM-CSF Immunostimulant fusion protein.
 XX
 KW Her2-GM-CSF; granulocyte macrophage colony stimulating factor;
 KW growth factor receptor; oncogene; immunostimulant; cancer;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..653
 FT Peptide /Label= Her2
 FT /Label= 654..655
 FT /Label= Linker
 FT Protein 656..782
 FT /Label= GM-CSF
 XX
 PN WO9724438-A1.
 PD 10-JUL-1997.
 XX
 PF 23-DEC-1996; 96WO-US20241.
 XX
 PR 28-DEC-1995; 95US-0579823.
 XX
 PA (ACTI-) ACTIVATED CELL THERAPY INC.
 PI Laus R, Ruegg CL, Wu H;
 XX
 DR WPI; 1997-363674/33.
 DR N-PSDB; AAT72725.
 XX
 PT Potent APC that activates T-cells to give multivalent cellular
 PT immune response - can also induce a cytotoxic T-cell response in a
 PT vertebrate subject
 PS Disclosure; Fig 8; 45pp; English.
 CC
 CC A fusion protein (AAM19764) comprises Her2 (a growth factor receptor
 CC that is overexpressed in breast, ovarian and other cancer cells)
 CC and granulocyte-macrophage colony stimulating factor (GM-CSF). It
 CC is the expression product of a nucleic acid molecule (AAT72725)
 CC prep. by PCR amplification of Her2 cDNA from a breast cancer cell
 CC line and fusion to GM-CSF cDNA. Fusion expression vectors can be
 CC used to transfect mammalian and insect cells. The Her2-GM-CSF
 CC fusion protein is used to generate anti-Her2 immunity. Tumour
 CC cells are eliminated by cytotoxic T lymphocytes activated in vivo
 CC or in vitro by exposure to antigen-presenting cells exposed to the
 CC fusion protein.
 CC
 CC
 SO Sequence 782 AA:
 Query Match 42.3%; Score 292; DB 18; Length 782;
 Best Local Similarity 100.0%; Pred. No. 4.7e-259;
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 38 GAASTOYCTGDMKRLRPASPEHLDMLRHLYGCGVVGNGLEETVPTNASLSFLDIO 97
 DB 19 GAASTOYCTGDMKRLRPASPEHLDMLRHLYGCGVVGNGLEETVPTNASLSFLDIO 78
 QY 98 EVGGYVLIANNOVROVPLQRLIRVGTQLFEDNVALAVLNDGDP LNNTPTVGTASPGGLR 157
 DB 79 EVGGYVLIANNOVROVPLQRLIRVGTQLFEDNVALAVLNDGDP LNNTPTVGTASPGGLR 138
 QY 158 ELQIRSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPM 217
 DB 139 ELQIRSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPCSPM 198
 QY 218 CKGSRGCGESSEDCOSLTRIVCAGGCARCKGRLPTDCCHQCAAGCGPRHSDCLACLHF 277
 DB 199 CKGSRGCGESSEDCOSLTRIVCAGGCARCKGRLPTDCCHQCAAGCGPRHSDCLACLHF 258

OY 278 NMSGICELCPALVTYNTDTFESMPNEGRTYFGASCYACPYNYLSTDVGS 329
 |||||
 DB 259 NMSGICELCPALVTYNTDTFESMPNEGRTYFGASCYACPYNYLSTDVGS 310

RESULT 13 AAB21203

ID AAB21203 standard; protein; 919 AA.

AC AAB21203;

DT 12-JAN-2001 (first entry)

DE Human HER-2/neu fusion protein.

KM Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
 breast cancer; prostate cancer; ovarian cancer; lung cancer;
 colon cancer; fusion protein.

OS Homo sapiens.
 Synthetic.

PN WO200044899-A1.

PD 03-AUG-2000.

PF 28-JAN-2000; 2000WO-US02164.

PR 29-JAN-1999; 990S-0117976.

PA (CORI-) CORIXA CORP.
 (SMIK) SMITHKLINE BEECHAM.

PI Cheever MA, Gheysen D;

DR WPI: 2000-505976/45.

PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
 useful for vaccinating against breast, ovarian, colon, lung and
 prostate cancers -

PS Claim 2; Fig 12; 128bp; English.

CC The present sequence is a fusion protein comprising the extracellular
 domain and the phosphorylation domain of the human HER-2/neu protein.
 CC HER-2/neu is a member of the tyrosine kinase family of receptor-like
 CC glycoproteins and shows homology to the epidermal growth factor receptor
 CC (EGFR). It probably plays a part in cell growth and/or differentiation.
 CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used
 CC to treat or prevent cancer by eliciting or enhancing an immune response
 CC to the HER-2/neu protein. They may be used to treat malignancies such as
 CC breast, ovarian, colon, lung and prostate cancers, and may be used as an
 CC antigen to vaccinate against these neoplasias.

CC Sequence 919 AA;

Query Match 42.3%; Score 292; DB 21; Length 919;

Best Local Similarity 100.0%; Pred. No. 5.4e-259;

Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 38 GAASTOVCTGTDMLKRLPASPETHLMLRLHYGCGVOGNTLFTVPTNASLSTFODIO 97
 |||||
 DB 19 GAASTOVCTGTDMLKRLPASPETHLMLRLHYGCGVOGNTLFTVPTNASLSTFODIO 78
 |||||
 OY 98 EVQGVYLIANNOYRQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLNNTPTVTCASPGGLR 157
 |||||
 DB 79 EVQGVYLIANNOYRQVPLQRLIRVGTQLEFEDNYALAVLDNGDPLNNTPTVTCASPGGLR 138
 |||||
 OY 158 ELQRLSTELTKGVLQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPM 217
 |||||
 DB 139 ELQRLSTELTKGVLQIRNPOLCYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPM 198
 |||||

OY 218 CKGSRGWGESSBDCSLRTFTVCAGCARCKGPLPTDCCHQDCAAGCTGKHSDCIACLHF 277
 |||||
 DB 199 CKGSRGWGESSBDCSLRTFTVCAGCARCKGPLPTDCCHQDCAAGCTGKHSDCIACLHF 258
 |||||
 OY 278 NMSGICELCPALVTYNTDTFESMPNEGRTYFGASCYACPYNYLSTDVGS 329
 |||||
 DB 259 NMSGICELCPALVTYNTDTFESMPNEGRTYFGASCYACPYNYLSTDVGS 310

RESULT 14 AAM51148

ID AAM51148 standard; protein; 919 AA.

AC AAM51148;

DT 17-JUN-2002 (first entry)

DE Her-2/neu extracellular domain-phosphorylation domain fusion.

KM Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 tyrosine kinase; receptor; c-erbB2; gene therapy.

OS Homo sapiens.

PN WO200212341-A2.

PD 14-FEB-2002.

PF 03-AUG-2001; 2001WO-US24283.

PR 03-AUG-2000; 2000US-0632507.

PA (CORI-) CORIXA CORP.
 (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Cheever MA, Gheysen D;

PT Her-2/neu fusion protein for treating or preventing cancer by eliciting
 or enhancing an immune response to the protein, has Her-2/neu
 extracellular domain fused to Her-2/neu intracellular or
 phosphorylation domain -

PS Claim 2; Fig 12; 141bp; English.

CC The present sequence is that of a fusion protein between the
 CC extracellular domain and phosphorylation domain of human Her-2/neu
 CC (see AAM51143), an oncogenic self-protein and target for anti-cancer
 CC vaccines. The fusion protein can be obtained by recombinant DNA
 CC methods. Her-2/neu overexpression correlates with a poor prognosis
 CC in breast and ovarian cancers. The invention provides Her-2/neu
 CC fusion proteins, nucleic acids encoding them, viral vectors, and
 CC vaccines comprising the fusion proteins or nucleic acid molecules.
 CC In preferred fusion proteins, the extracellular domain of a
 CC Her-2/neu protein is fused to a Her-2/neu intracellular domain or
 CC phosphorylation domain (or its delta-pd fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal *ex vivo* with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and iso-ated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.

```
XX Sequence 919 AA:
SQ
Query Match 42.3%; Score 292; DB 23; Length 919;
Best Local Similarity 100.0%; Pred. No. 5.4e-259;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASTOVCTGDMKRLRPASPETHLDMRLHYOCQVVGMLLELYLPTNASLSFLDIO 97
DB 19 GAASTOVCTGDMKRLRPASPETHLDMRLHYOCQVVGMLLELYLPTNASLSFLDIO 78
QY 98 EVQGVLLAHNQVQVPLQRLRIYRGVQLFEDNTALAVLDGDDPLNNTPTVTGASPGGLR 157
DB 79 EVQGVLLAHNQVQVPLQRLRIYRGVQLFEDNTALAVLDGDDPLNNTPTVTGASPGGLR 138
QY 158 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
QY 218 CKGRMCWESSEDCOSLTRVACAGCARGKPLPTDCHEOCAGCTGPKHSDCLACLHF 277
DB 199 CKGRMCWESSEDCOSLTRVACAGCARGKPLPTDCHEOCAGCTGPKHSDCLACLHF 258
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 310

RESULT 15
AAB21208
ID AAB21208 standard; Protein: 1200 AA.
XX
AC AAB21208;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human HER-2/neu protein.
XX
KW Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
KW colon cancer.
XX
OS Homo sapiens.
XX
PN M0200044899-A1.
XX
PD 03-AUG-2000.
XX
PE 28-JAN-2000; 2000MO-US02164.
XX
PR 29-JAN-1999; 99US-0117976.
XX
PA (CORI-) CORIXA CORP.
PA (SMIK) SMITHKLINE BEECHAM.
XX
PI Cheever MA, Gheysen D;
XX
XX WPI: 2000-505976/45.
XX DR N-PSDB: AAA89736.
XX
PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
PS Disclosure: Fig 15; 128bp; English.
XX
CC The present sequence is the human HER-2/neu protein. It is a member
CC of the tyrosine kinase family of receptor-like glycoproteins and shows
CC homology to the epidermal growth factor receptor (EGFR). It probably
CC plays a part in cell growth and/or differentiation. The HER-2/neu
CC gene is an oncogene. An HER-2/neu fusion protein comprising a
CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
CC domain may be used to treat or prevent cancer by eliciting or
```

```
CC enhancing an immune response to the HER-2/neu protein. It may be used
CC to treat malignancies such as breast, ovarian, colon, lung and
CC prostate cancers, and may be used as an antigen to vaccinate against
CC these neoplasias.
XX
SQ Sequence 1200 AA:
Query Match 42.3%; Score 292; DB 21; Length 1200;
Best Local Similarity 100.0%; Pred. No. 6.8e-259;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GAASTOVCTGDMKRLRPASPETHLDMRLHYOCQVVGMLLELYLPTNASLSFLDIO 97
DB 19 GAASTOVCTGDMKRLRPASPETHLDMRLHYOCQVVGMLLELYLPTNASLSFLDIO 78
QY 98 EVQGVLLAHNQVQVPLQRLRIYRGVQLFEDNTALAVLDGDDPLNNTPTVTGASPGGLR 157
DB 79 EVQGVLLAHNQVQVPLQRLRIYRGVQLFEDNTALAVLDGDDPLNNTPTVTGASPGGLR 138
QY 158 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 217
DB 139 ELQRLSLTEILKGGVLIQIRNPOLCYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPM 198
QY 218 CKGRMCWESSEDCOSLTRVACAGCARGKPLPTDCHEOCAGCTGPKHSDCLACLHF 277
DB 199 CKGRMCWESSEDCOSLTRVACAGCARGKPLPTDCHEOCAGCTGPKHSDCLACLHF 258
QY 278 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 329
DB 259 NMSGICELHCPALVTYNTDFESMPNPEGRTYFGASCTACPYNYLSTDVGS 310
```

Search completed: April 28, 2003, 13:44:37
Job time : 43 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 8, 2003, 12:36:28 ; Search time 210 Seconds
(Without alignments)
4083.173 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAPILLARASLSIGFLF.....EPYQEGAPRRPAHHNNHH 690

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=rlp
-Q=/cgn2_1/USFTO.spool/US09821883/runatc_28042003_104750_14608/app_query.fasta.1.839
-DB=published.Applications.NA-QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.ccd -LIST=45 -DOCCALIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MALIGN=200000000 -USER=US09821883.ecgn.1.1.80.etrnatc_28042003_104750_14608
-NCU=6 -ICPU=3 -NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOS
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3739	100.0	2070	US-09-821-883-7	Sequence 7, Appl1
2	3473.5	92.9	2091	US-09-821-883-9	Sequence 9, Appl1
3	2974	79.5	1665	US-09-821-883-6	Sequence 6, Appl1
4	2959.5	79.2	1692	US-09-821-883-8	Sequence 8, Appl1

5	2426	64.9	4543	9	US-09-769-508-1	Sequence 1, Appl1
6	2423	64.8	4530	9	US-09-877-177-11	Sequence 11, Appl1
7	2422	64.8	9274	10	US-09-811-123-7	Sequence 7, Appl1
8	2422	64.8	9274	10	US-09-811-115-1	Sequence 1, Appl1
9	2418	64.7	4473	9	US-09-441-411-5	Sequence 5, Appl1
10	2410.5	64.5	3768	10	US-09-811-123-8	Sequence 8, Appl1
11	2410.5	64.5	3768	10	US-09-811-115-2	Sequence 2, Appl1
12	2405.5	64.3	3768	9	US-09-854-356-9	Sequence 9, Appl1
13	2405.5	64.3	3768	9	US-09-930-125-1	Sequence 1, Appl1
14	2188.5	58.5	1437	10	US-09-821-883-10	Sequence 10, Appl1
15	1968	52.6	3935	9	US-09-854-356-10	Sequence 10, Appl1
16	1968	52.6	3935	9	US-09-870-759-117	Sequence 117, Appl1
17	1961	52.4	3771	9	US-09-854-356-11	Sequence 11, Appl1
18	1587	42.4	867	10	US-09-821-883-24	Sequence 24, Appl1
19	1199.5	32.1	1191	10	US-09-821-883-28	Sequence 28, Appl1
20	1188	31.8	1806	9	US-09-930-125-5	Sequence 5, Appl1
21	1183	31.6	1755	9	US-09-930-125-6	Sequence 6, Appl1
22	1183	31.6	1767	9	US-09-930-125-4	Sequence 4, Appl1
23	1183	31.6	1773	9	US-09-930-125-7	Sequence 7, Appl1
24	1182	31.6	651	10	US-09-821-883-26	Sequence 26, Appl1
25	1182	31.6	3537	10	US-09-821-883-30	Sequence 30, Appl1
26	803	21.5	1115	9	US-10-102-806-165	Sequence 165, Appl1
27	724.5	19.4	2643	9	US-10-172-620-15	Sequence 15, Appl1
28	724.5	19.4	3633	10	US-09-725-433-1	Sequence 1, Appl1
29	724.5	19.4	10058	9	US-09-974-298-98	Sequence 98, Appl1
30	721	19.3	5484	10	US-09-920-300A-1731	Sequence 1731, Appl1
31	720.5	19.3	5264	10	US-09-920-300A-1731	Sequence 1731, Appl1
32	720.5	19.3	5264	12	US-10-033-528-1731	Sequence 1731, Appl1
33	714.5	19.1	4879	9	US-10-172-620-13	Sequence 13, Appl1
34	714.5	19.1	4975	10	US-09-880-107-2342	Sequence 2342, Appl1
35	713.5	19.1	2601	10	US-09-940-101-3	Sequence 3, Appl1
36	710	19.0	1958	10	US-09-867-521-1	Sequence 1, Appl1
37	684	18.3	1833	10	US-09-783-708-2	Sequence 2, Appl1
38	675	18.1	1011	12	US-10-044-090-509	Sequence 509, Appl1
39	675	18.1	1318	9	US-10-228-811-3	Sequence 3, Appl1
40	674	18.0	435	9	US-09-826-025-8	Sequence 8, Appl1
41	674	18.0	435	9	US-10-083-590-14	Sequence 14, Appl1
42	669	17.9	452	9	US-09-918-995-12765	Sequence 32765, Appl1
43	668	17.9	381	10	US-09-821-883-19	Sequence 19, Appl1
44	655.5	17.5	1046	10	US-09-925-301-390	Sequence 390, Appl1
45	610	16.3	731	12	US-10-044-090-826	Sequence 826, Appl1

ALIGNMENTS

RESULT 1
US-09-821-883-7
: Sequence 7, Application US/09821883
: Patent NO. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Rehner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022-30
CURRENT APPLICATION NUMBER: US/09-821-883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 2070
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500-HGM-CSF construct
US-09-821-883-7
Alignment Scores: 3.36e-315 Length: 2070
Pred. No.: 3739.00 Matches: 690
Score:


```
RESULT 2
US-09-821-883-9
: Sequence 9, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Relner
: APPLICANT: Vidovic, Damir
: APPLICANT: Gradis, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE OF INVENTION: Cell-Based Immunotherapy
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 2091
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500*-rGM-CSF construct
: US-09-821-883-9

Alignment Scores:
Pred. No.: 3,66e-292 Length: 2091
Score: 3473.50 Matches: 641
Percent Similarity: 94.42% Conservative: 19
Best Local Similarity: 91.70% Mismatches: 28
Query Match: 92.90% Indels: 11
DB: 10 Gaps: 2

US-09-821-883-2 (1-690) x US-09-821-883-9 (1-2091)
QY 1 MetArGAlaAlaProLeuLeuLeuAlaArGAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGACCCCTCTCTCTGCGCCAGGCGAGAACCTTAGCTTGCTTCTTGT 60
QY 21 LeuLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 40
Db 61 CTGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 41 SerThrGlnValCysThrGlnValCysThrGlnValCysThrGlnValCysThrGlnVal 60
Db 121 TCGACCCAGAGTGTGACCGGACAGACATGAACTGCGGCTCCCTGCGCAGTCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuArgHisLeuArgHisLeuArgHisLeuArgHisLeu 80
Db 181 CACCTGGAGCATGCTCCGCGCCTCTACACAGGCGCTCCGCGCGTGTGCGAGGAACCTGGA 240
QY 81 LeuThrTyLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTCACCTACCTGCGCCACCAATGCGCAGCTCTCTCTGCGAGATATCAGAGAGTGCAG 300
QY 101 GlyTyValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GCGTACGCTGCTATGCTGCTACACCAAGTGAAGGCGTCCACCTGCAAGAGCTGCGGATT 360
QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGGAGAGCACCCAGCTCTTTGAGGACAACTATGCTCGGCGCTGCTAGCAATGGAAC 420
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACAATACACCCCTGTGCAAGAGGCGCTCCCGAGAGAGGCTCGGAGAGCTGAG 480
QY 161 LeuArgSerLeuThrGlnIleLeuLeuGlyGlyValLeuIleGlnArgAsnProGlnIleu 180
Db 481 CTTGAGACCTTCACAGAAATCTTGAAAGAGGAGGCTTGATTCACCGGAGAACCCCAAGCTC 540
QY 181 CysTyArgIleAspThrIleLeuThrPheAspIlePheHisIleAsnGlnLeuAlaLeu 200
```

```
Db 541 TGTACACGAGACAGATTTTGTGAGAGACATCTTCCACAGAAACCAACCTGGCTCTC 600
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysGly 220
Db 601 ACACGTATAGACACCAACCGCTCTGCGGCTGCGACCCCTGCTCTCCATGTGAAGGC 660
QY 221 SerArgCysThrProGlyGlnSerSerGlnAspCysGlnSerIleLeuThrArgThrValCysAla 240
Db 661 TCCGCTGCTGGGAGAGATTTGTGAGATTTGTGAGACCTGAGCGCGACGTGTCTGTC 720
QY 241 GlyGlyCysAlaArgCysValGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
Db 721 GGTGGCTGTGCGCCCTGCAAGGGCGACCTGCGCAGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 261 AlaGlyCysThrGlyProLeuHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCGGCTGTCAGGGCGCCCAACACTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 281 GlyTleGysGluLeuHisCysProAlaLeuValThrTyAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 301 MetProAsnProGluGlyArgTyThrPheGlyAlaSerCysValThrAlaCysProTy 320
Db 901 ATGCCCATCCCGAGGGCGCGGTATACATTCGCGCGCAGCTGTGTGACTGTGCTGTAC 960
QY 321 AsnTyLeuSerThrAspValGlySer-----GlyAla 331
Db 961 AACTACCTTTCTAGGAGAGCTGAGATCCCGCTGATCATTAATTCGAGAACGTTGGCCCT 1020
QY 332 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyValAspLeu 351
Db 1021 GGGGCTATGTCACACAGGACGCGACGCTCATCTACAGAGTGGGGTGGGAGCTTG 1080
QY 352 ThrLeuGlyLeuGluProSerGluGluValAlaProArgSerProLeuAlaProSerGlu 371
Db 1081 ACACTAGGCTGAGAGCCCTGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
QY 372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaValGlyGlyLeuGln 391
Db 1141 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 392 SerLeuProThrHisAspProSerProLeuGlnArgTySerGlnAspProThrValPro 411
Db 1201 AGCTTCCCAACATGACCCCAAGCCCTTACAGCGGTACAGTGAAGACCCCAACACTACC 1260
QY 412 LeuProSerGluThrAspGlyTyValAlaProLeuThrCysSerProGlnProGluTy 431
Db 1261 CTGCCCTGTGAGACTGATGGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 432 ValAsnGlnProAspValArgProGlnProProSerProAlaGlyGlyProLeuProAla 451
Db 1321 GTGAACCAAGCAGATGTTGCGGCGCCAGCCCTTGCAGGAGAGGCGCTGCTGCTGCT 1380
QY 452 AlaArgProAlaGlyAlaThrLeuGlnArgAlaValThrLeuSerProGlyValAsnGly 471
Db 1381 GCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 472 ValValIleAspValPheAlaPheGlyGlyValAlaValGluAsnProGluTyLeuThrPro 491
Db 1441 GTGCTGAAGAAGCTTTTGGCTTTGGGGGTGCTGCGGAGAACCCCGACTTGTGACACC 1500
QY 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511
Db 1501 CAGGAGAGAGTGGCGCCACACCCCACTCTGCTGCTTCAAGCCCTTGTGACAAAC 1560
QY 512 LeuTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 531
Db 1561 CTTATTTACTGGACACAGACCCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
QY 532 ThrProThrAlaGluAsnProGluTyLeuGlyLeuAspValProAlaAlaAlaProAla 551
```


Db 1621 ACACCTACGCGAGAACCCAGACTACCTGGCTGAGCTGCCAGCGGCCGCCACC 1680
Qy ArgSerProSerProSerThrGlnProThrPGLNHISValAsnAlaIleGlnIuAlaArg 571
Db 1681 CGCTCAACCCAAACCTGTCACCGGCTGAGACATGTATGCAATCAAGAAAGCTCTG 1740
Qy 572 ArgLeuAsnLeuSerArgAspThrAlaIleGluMetAsnGluThrValGluValIle 591
Db 1741 AGCCTCTTAATGACATGCTGCTGTGAGAACGAAAAGAACAGACATGACATCATC 1800
Qy 592 SerGluMetPheAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluIleThrVal 611
Db 1801 TCATAAGAGTTCCTCCATCCAGAGCCGACATGTGTGACAGCCGCTGAGGATTAACAG 1860
Qy 612 GlnGlyLeuArgGlySerLeuThrIleValLeuGlyProLeuThrMetAlaSerHis 631
Db 1861 CAGGCTTACGGGCAACCTCAACCAATGATGCGCCCTTGAACATGATAGCAGCAC 1920
Qy 632 TyrLeuGlnHisCysProProThrProGluThrSerCysAlaThrGlnIleIleThrPhe 651
Db 1921 TACACAGACGAACCTGCTCCACACCGGAAACGACATGATAATAGAACACACCTTT 1980
Qy 652 GluSerPheLeuGlyLeuAsnLeuLysAspPheLeuValIleProPheAspCysThrGlu 671
Db 1981 GAGGATTTCTATAAGAACCTTAAGGCTTCTGTGATATCCCTTTTGACTGTGAGAG 2040
Qy 672 ProValGlnGluIuAlaProProProProAlaAlaHisHisHisHisHis 690
Db 2041 CCGCTCCAGAAAGGCGGCGCACCCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2091

RESULT 3
US-09-821-883-6
Sequence 6, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovic, Damir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821, 883
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193, 504
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6

LENGTH: 1665
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500 construct
US-09-821-883-6

Alignment Scores:
Pred. No.: 5,94e-249 Length: 1665
Score: 2974.00 Matches: 554
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 1
Query Match: 79.54% Indels: 3
DB: 10 Gaps: 1

US-09-821-883-2 (1-690) x US-09-821-883-6 (1-1665)

Qy 1 MetArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGACCCCTCCCTGCGCAGGCGCAAGGCTTGTGCTTCTTGTGCTT 60
Qy 21 LeuLeuPhePheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
Db 61 CTGCTTTTTCGCTGCTAGACCAAGTGTACTAGCCAGAGAGTGGCGCGCGCGCGCG 120

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCAAATGTGACACCGGACAGACATGAACTGGCGCTCCCTGCCACTCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlnValAsnLeuGlu 80
Db 181 CACCTGACATGCTCCCGCACCTCTTACAGAGGCTGCGAGGTGGCGGAGAAACCTGGAA 240
Qy 81 LeuThrThrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIuValGln 100
Db 241 CTCACCTACCTGGCCACCAATGCGACGCTGCTCTCTGAGATATCCAGAGAGTGCAG 300
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnThrLeuArgIle 120
Db 301 GCGTACCTGCTCATCCGCTCAACCAAGTGGAGAGTCCCATGCGAGCGCGGATTT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAspThrValAlaLeuValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGACCCAGCTCTTGTAGAGACATGACCTGGCGCTGCTAGCAATGAGAGAC 420
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACAATACCAACCTGTCACAGGGGCGCTCCCGAGAGGCGTGGGAGCTGCAG 480
Qy 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTCGAAGCCTCACAGAGATCTTGAAAGGAGGGGTCTGTGATCCAGCGGAACCCCGACCTC 540
Qy 181 CysTyrGlnAspThrIleLeuThrPheLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
Db 541 TGCTACAGGACACGATTTGTGTGAGAGACATCTTCACAGAACACCACTGCTGCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACACTGATACACCAACCAACGCTCTCGGCGCTGCACACCGCTTCTCGAGATGTAAAGGCG 660
Qy 221 SerArgCysThrPheLysLeuSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGGCTGCGGGAGAGATTTGTGAGATTTGTGAGAGCTGACGCGCTGCTGCTGCTG 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260
Db 721 GGTGGCTGTCCCGCTGCAGAGGGCCACTGCCACTGACTGCTGCTGCTGCTGCTGCTGCT 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCGGGGCGCAGGGGCCCAAGCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 281 GlyTyrLeuGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGCTGCACTGCCAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProThr 320
Db 901 ATGCCCAATCCGAGGCGCGGTATACATTTGGCGCCAGCTGTGTACTGCTGCTGCTGCT 960
Qy 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyIleValHisHisArgHisArg 340
Db 961 AACTACCTTTCTAGGACGCTGGATCGGGCGCTGGGGGATGATGCCACACAGCAGCAGCG 1020
Qy 341 SerSerSerThrArgSerGlyGlyAlaSerPheLeuThrLeuGlyLeuProSerGluGlu 360
Db 1021 ACCTCATCTACACGAGGTGGCTGGGAGACTACACTAGAGGCTGGGCGCTCTGAAAG 1080
Qy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
Db 1081 GAGGCCCCCAGGCTCTCACTGCGACCTCCGAAGGGGCTGGCTCCGATGTATTTGATGCT 1140
Qy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
Db 1141 GACCTGGGAATGGGGGAGCAGCAAGGGGCTGCAAGAGCTCCCAACATGACCCAGCGCT 1200
Qy 401 LeuGlnArgTyrSerLysAspProThrValProLeuProSerGluThrAspGlyThrVal 420

Db 1201 CTACAGCGGTACGTGAGAGACCCACACTACCCCTGAGACGATGAGCTACGTT 1260
 Oy 421 AAlaProLeuThrCysSerProGlnProGluTyrValAlaGlnIProAspValArgProGln 440
 Db 1261 GCCCGCCCTGACCTGAGCCCGCCAGCTTAATATGTGAACACGCCGATGTTGGGCCAG 1320
 Oy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
 Db 1321 CCCCCCTTGGCCCCGAGAGGGCCCTGCTGCTGCTGCCACCTGCTGGGCCACTTGGAA 1380
 Oy 461 ArgAlaIleThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
 Db 1381 AGGCGCAAGACTCTCTCCCGAGGAGAAATGGGGCTGCAAAAGAGCTTTGGCTTGGG 1440
 Oy 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHis 500
 Db 1441 GGTGCGGTGAGAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCCCTACGCCAC 1500
 Oy 501 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 520
 Db 1501 CCTCTCTGCTGCTTACGCGCAGCCCTTGACAACTCTATTACTGGAGACAGACCCACCA 1560
 Oy 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540
 Db 1561 GAGCGGGGGCTCCACCCAGCACCTTCAAGGAGACCTAGGCGAGAGAACCCAGAGTAC 1620
 Oy 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSer 557
 Db 1621 CTGGGTCTGAGCTGACGAGCGCC-----GCACA-TCACCATCCACATCA 1664

RESULT 4

US-09-821-883-8

Sequence 8, Application US/09821883

Patent No. US20020061310A1

GENERAL INFORMATION:

APPLICANT: Laus, Relner

APPLICANT: Vidovic, Damir

APPLICANT: Gradalis, Thomas

TITLE OF INVENTION: Compositions and Methods for Dendritic

FILE REFERENCE: 7636-0022.30

CURRENT APPLICATION NUMBER: US/09/821,883

CURRENT FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: US 60/193,504

PRIOR FILING DATE: 2000-03-30

NUMBER OF SEQ ID NOS: 30

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 1692

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: HER500* construct

US-09-821-883-8

Alignment Scores:

Pred. No.: 1,1e-247 Length: 1692
 Score: 2959.50 Matches: 554
 Percent Similarity: 97.88% Conservative: 0
 Best Local Similarity: 97.88% Mismatches: 1
 Query Match: 79.15% Indels: 12
 DB: 10 Gaps: 2

US-09-821-883-2 (1-690) x US-09-821-883-8 (1-1692)

Oy 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
 Db 1 ATGAGAGCTGACCCCTCTCTGCGCAGGCGAGCAAGCCTTAGCCTTCTTGT 60
 Oy 21 LeuLeuPhePheThrPheLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAla 40
 Db 61 CTGCTTTTCTGCTGAGACCGAAGTGTACTAGCCAAAGAGATTGGCGCGGCCCGC 120

Oy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuAlaGluProAlaSerProGluThr 60
 Db 121 TCGACCAAGTGTGACCGGCGACAGACATGAAGCTGGGGCTCCCTCCAGTCCGAGAC 180
 Oy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnLysLeuGlu 80
 Db 181 CACCTGAGCATGTCTCCGACCTTACACAGGGCTCCAGTGTGTCAAGGAACCTGGAA 240
 Oy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleValGln 100
 Db 241 CTCACCTACCTGCGCCACCAATGCCAGCTGCTTCGACAGATATCCAGAGGTGAG 300
 Oy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
 Db 301 GGCTACGTGCTCATGTGCTCACACCAAGTGAAGGCTCCACTGACAGAGTGGGAT 360
 Oy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
 Db 361 GTGCGAGGACCCACCTCTTGTAGGACAAATGCTGCGCTGCTAGACAAATGAGAC 420
 Oy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
 Db 421 CCGTGAAACAATACCCCTGTACAGGGGCTCCCGAGAGGCTCGGGAGCTGCAG 480
 Oy 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
 Db 481 CTTCGAAAGCTTCACAGAGATCTTGAAGAGGGGCTTGATGCACGGAACCCAGCTGC 540
 Oy 181 CysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
 Db 541 TGCTACCAAGACACGATTTTGTGAAGACATCTTCCACAGAACACAGCTGCTCTGC 600
 Oy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
 Db 601 ACAGTATAGACACCAACCGCTCTGCGCTGCACCCCTGTTCTCCGATGTGAAGGC 660
 Oy 221 SerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValLysAla 240
 Db 661 TCCGCTCTGCGGAGAGAGTCTGAGATGTCTGAGCGTCAAGCGTCAAGCGCTGCTG 720
 Oy 241 GlyLysCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluLincysAla 260
 Db 721 GGTGCTGTGCGCTGCAAGGGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Oy 261 AlaGlyCysThrGlyProLysHisSerAspLysAlaCysLeuHisPheAsnHisSer 280
 Db 781 GCGGCTGCGAGGCGCCCAAGACATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Oy 281 GlyIleCysGluLeuHisCysProAlaLeuValIleThrTyrAsnThrAspThrPheGluSer 300
 Db 841 GGCATCTGTGAGCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 Oy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValIleThrLacysProTyr 320
 Db 901 ATGCCCAATCCCGAGGCGCGGTATGATGCTGCGCGCTGCTGCTGCTGCTGCTGCT 960
 Oy 321 AsnTyrLeuSerThrAspValGlySer-----GlyAla 331
 Db 961 AACTACCTTTCTACGAGCTGGGATCCCGTACCATCATTAATTTCGAGAAAGTTGGCGCT 1020
 Oy 332 GlyLysMetValHisHisArgHisArgSerSerThrArgSerGlyLysLysLeu 351
 Db 1021 GGGGCGATGTGTCACACAGGACCGGACGCTATCTACACAGAGTGGGCGTGGGACGCTG 1080
 Oy 352 ThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 371
 Db 1081 ACACTAGGGCTGAGCCCTCTGTAAGAGAGGCGCCAGGTCTCCACTGCGACCTCCGAA 1140
 Oy 372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 391
 Db 1141 GGGGCTGCTCCGATGTATTGTATGTGACCTGGGAATGGGGGACCAAGGGGCTGCA 1200

```

OY 392 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 411
Db 1201 AGCCCTCCACACGATGACCCACCTCTACAGGGTACGTAGAGACCCACACTACC 1260
OY 412 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 431
Db 1261 CTGGCTCTGAGACTGATGGCTACCTTCCCTGACCTGACGCCCCCAGCTGAAAT 1320
OY 432 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 451
Db 1321 GTGAACCCAGCCAGATGTTCCGCCCCAGCCCCCTTCGCCGAGAGAGGCCCTGCTCT 1380
OY 452 AlaArgProAlaGlyAlaThrLeuGlnArgAlaGlyThrLeuSerProGlyLysAsnGly 471
Db 1381 GCCCGACCTGCTGGTGGCCACTCTGAAAGGGCCCAAGACTCTCTCCCGAGGAGATGG 1440
OY 472 ValValLysAspValPheAlaPheGlyGlyValAlaGluAsnProGluTyrLeuThrPro 491
Db 1441 GTCGTCAAGACGTTTGGCTTTGGGGGTCCCTGGAGAACCCCGAGTACTTGACACCC 1500
OY 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511
Db 1501 CAGGAGAGAGCTGCCCTCACCCCTCCTCTGCTGACGCCAGCCTTGACAC 1560
OY 512 LeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly 531
Db 1561 CTCTATTACTGGGACCCAGGACCCACAGAGGGGGCTCCACCCAGCACCTTCAAAAGG 1620
OY 532 ThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAla 551
Db 1621 ACACCTAGGAGAGAACCCAGACTACTGGGTCTGGACGTGCCAGCGCC-----GCA 1674
OY 552 ArgSerProSerProSer 557
Db 1675 CA-TCACCATCACCATCA 1691

RESULT 5
US-09-769-508-1
: Sequence 1, Application US/09769508
: Patent No. US2002015527A1
: GENERAL INFORMATION:
: APPLICANT: STUART, SUSAN G.
: APPLICANT: MONAHAN, JOHN J.
: APPLICANT: LANGTON, BEATRICE CLAUDIA
: APPLICANT: HANCOCK, MIRIAM E.C.
: APPLICANT: CHAO, LORRINE A.
: APPLICANT: BLUFORD, PETER
: TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
: FILE REFERENCE: BEBIO-111-C1
: CURRENT APPLICATION NUMBER: US/09/769,508
: CURRENT FILING DATE: 2001-01-26
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4543
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (150)..(3914)
US-09-769-508-1

Alignment Scores:
Pred. No.: 7, 1e-201
Score: 2426.00
Percent Similarity: 40.66%
Best Local Similarity: 40.51%
Query Match: 64.88%
DB: 9
Gaps: 4
US-09-821-883-2 (1-690) x US-09-769-508-1 (1-4543)
OY 2 ArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPhe----- 18

```

```

Db 63 CGGCCCCACCCCTTCGACAGACCCCGCGCCCCCTCCAGCCGGTCCAGCCGAGC 122
OY 19 -----LeuPheLeuPhePheThrLeuAsp 27
Db 123 CATGGGGCCGAGCCGAGTGAACATGAGAGCTGGGGGCTTGTGCCCTGG----- 176
OY 28 ArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThrGlnValCysThrGly 47
Db 177 ---GGGCTCTCTCCGCCCTTGGCCCCCGAGGCCGAGACCCCAAGTAGTGTGACCGCG 233
OY 48 ThrAspMetLysLeuArgLeuProLalaSerProGluThrHisLeuAspMetLeuArgHis 67
Db 234 ACAGACATGAAGCTGGGCTCTCCAGTCCCGAGCCACCTGGACATGCTCCGCCAC 293
OY 68 LeuTyrGlnGlyCysGlnValValGlnLysAsnLeuGluLeuThrTyrLeuProThrAsn 87
Db 294 CTCTACCGAGGCTGCGAGGTGGTGCAGGAAACCTGGAACTCCTACCTGCCACCAAT 353
OY 88 AlaSerLeuSerPheLeuGlnAspIleGlnLysValGlnGlyTyrValLeuIleAlaHis 107
Db 354 GCCAGCCTGTCTCTTCAGAGATATCCAGAGGTGAGGGCTAGCTCATGCTCATC 413
OY 108 AsnGlnValArgGlnValProLeuGlnArgLeuArgIleValArgGlyThrGlnLeuPhe 127
Db 414 AACCAAGTGAAGGAGGTCCACTCGAGAGCTGGGGATTGGCGAGGACCCACTCTTT 473
OY 128 GluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeuAspAsnThrThrPro 147
Db 474 GAGGACAACTATGCCCTGGCCGCTGCTAGACAAATGAGACCCGCTGACAAATACCACT 533
OY 148 ValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIle 167
Db 534 GTCAAGAGGGCTCCCGCCAGAGGCTGGGGAGGTGAGCTTCGAAGCCTCAGAGAGATC 593
OY 168 LeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeuCysTyrGlnAspThrIleLeu 187
Db 594 TTGAAGAGGGGTCTTGATCCAGGGGAACCCAGGCTCTCTACAGGACAGCATTTTG 653
OY 188 TrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArg 207
Db 654 TGAAGAGCATCTTCCACAAAGAACACACAGCTGGCTGCACATGATGACACCAACGCC 713
OY 208 SerArgAlaCysHisProCysSerProMetCysLysGlySerArgCysTrpGlyLysIle 227
Db 714 TCTGGGGCTGCCACCCCTGTTCTCCGATGTGAAGGGCTCCCGCTGGGGAGAGAGT 773
OY 228 SerGluAspCysGlnSerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLys 247
Db 774 TCTAGAGATTGTCAAGGCTGACGCGGCACTGTGTGCGGTGGCTGTGGCGCTGCAG 833
OY 248 GlyProLeuProThrAspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLys 267
Db 834 GGGCCAGTGGCCACTGACTGTGTCATGAGAGAGAGTGTGCGGGCTGCAGGGGCCAG 893
OY 268 HisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGlyIleCysGlnLeuHisCys 287
Db 894 CACTGTGACTCCCTGGCTGCTCCACTTCAACCAAGTGGCATGTGAGTGCCTCACTGC 953
OY 288 ProAlaLeuValThrTyrAsnThrAspThrPheGluSerMetProAsnProGluLysArg 307
Db 954 CCAGCCCTGGTCACTCAACACAGACACGTTTAGTCCATGCCCCAATGCCGAGGGCCGG 1013
OY 308 TyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspVal 327
Db 1014 TATACATTTCGGCGGACACTGTGTACTGTGCTGCTCCATCAAACTACTTTCACGAGAGTG 1073
OY 328 GlySer----- 329
Db 1074 GGATCTTGACACCTGTCTGCCCTTCACAAACCAAGAGTGACAGAGAGATGAACA 1133
OY 329 ----- 329

```

Db 1134 CAGCGGTGTGAGAACTGCAGCAAGCCCTGTGCCAGTGTCTATGTCTGGCATGGAG 1193
Oy 329 ----- 329
Db 1194 CACTTGCAGAGGTGAGGGAGCTTACCAAGTCCAAATATCCAGAGATTGTGCTGGCTGCAAG 1253
Oy 329 ----- 329
Db 1254 AAGATCTTTGGAGGCTGGCAATTTCTGCCGGAGACTTTGATGGGGACCCAGCCTCCAAAC 1313
Oy 329 ----- 329
Db 1314 ACTGCCCGCTCCAGCCAGAGCAGCTCCAAAGTTTGTGAGACTGTGGAAGATCACAGGT 1373
Oy 329 ----- 329
Db 1374 TACCTATACATCTCAGCATGGCCGGACAGCCCTGACCTCAGCGTCTTCCAGAACTCG 1433
Oy 329 ----- 329
Db 1434 CAAGTATCCGGGGAGCAATTTCTGCACAATGGCGCTACTGCTGACCTGCAAGGGCTG 1493
Oy 329 ----- 329
Db 1494 GGCATCAGCTGGCTGGGGCTGGCTCAGTGAAGGAATGGGCACTGGACTGGCTCAATC 1553
Oy 329 ----- 329
Db 1554 CACCAATACCCCACTCTGCTTCTGCTGCACAGGTGCCCTGGGACCACTCTTTGGGAAC 1613
Oy 329 ----- 329
Db 1614 CCGCACAAGTCTGTCTCCACACTGCCAAACCGGCCAGAGAGAGTGTGGCGAGGCG 1673
Oy 329 ----- 329
Db 1674 CTGGCTCCACACAGCTGTGCGCCCGAGGCACTGTGGGTCAAGGCCACCACTGT 1733
Oy 329 ----- 329
Db 1734 GTCAACTGCAGCAGTTCTTCTGGGGCCAGAGTGCCTGGAGGAATGCCAGTACTGCAG 1793
Oy 329 ----- 329
Db 1794 GGGCTCCCAAGGAGTATGTGAATGCAGGCACTGTTGGCCTGCACACCTGATGTGAG 1853
Oy 329 ----- 329
Db 1854 CCCCAGAAATGGCTCAGTAGCCTGTTTGGACCGAGGCTGACAGTGTGTGGCTGTGCG 1913
Oy 329 ----- 329
Db 1914 CACTAATAGGACCTCCCTTCTTGGCGTGGCCCGCTGCCAGCGGTGTGAACCTGACTTC 1973
Oy 329 ----- 329
Db 1974 TCCTATATGCCCATCTGGAAGTTTCCAGATGAGGAGGGGCATGGCAGCCTTGCCCATC 2033
Oy 329 ----- 329
Db 2034 AACTGACCCACTCTGTGTGTGACCTGTGATGACAAGGGCTGCCCGCGAGAGAGAGCC 2093
Oy 329 ----- 329
Db 2094 AGCCCTTGAGCCTCATGTCTGTGGGTGTGGCATTTGCTGTGCTGTGCTTTGGGG 2153
Oy 329 ----- 329
Db 2154 GTGTCTTTGGGATCCTCATCAAGGAGCGGAGAGCAAGAAATCCGAAGTACACGATGGCG 2213
Oy 329 ----- 329
Db 2214 AGACTGCTGCAAGAAACGAGCTGTGTGAGCGCTGACACTAGCGAGCATGGCCAAAC 2273

Oy 329 ----- 329
Db 2274 CAGCGCATGATCGGATCTCTGAAGAGACGAGCTGAGGAAGTCAAGCTGTGATCT 2333
Oy 329 ----- 329
Db 2334 GGGCCTTTTGGACAGTCTACAAAGGCACTGTGATCCCTGATGGGAGAAATGTGAAAT 2393
Oy 329 ----- 329
Db 2394 CCAGTGGCCATCAAGTGTGAGGAAAAACACATCCCCAAAGCCAAAGAAATCTTA 2453
Oy 329 ----- 329
Db 2454 GACGAAGCATACGTATGCTGTGTGGGCTCCCATATGCTCCCGCTTGTGGCATC 2513
Oy 329 ----- 329
Db 2514 TGCTGACATCCAGCTGACGTGTGTGACAGCTTATGCCCTATAGCTGCTTTAGAC 2573
Oy 329 ----- 329
Db 2574 CATGTCCGGAAAAACCGGAGCGCTGGGCTCCAGACCTGTGAACTGTGTATGCAAG 2633
Oy 329 ----- 329
Db 2634 ATTGCCAAGGGATGAGCTACTGAGGAGATGTGGCGCTGTACACAGGAGACTTGGCCCT 2693
Oy 329 ----- 329
Db 2694 CCGAAGTGTGTGTCAAGATGCCAACCATGTCAAAATTACAGACTTGGGCTGGCTGG 2753
Oy 329 ----- 329
Db 2754 CTGTGACATTGACAGAGACAGATGACATGAGTGGGGCAAGGTGCCCTCAAGTGG 2813
Oy 329 ----- 329
Db 2814 ATGGGCTGTGAGTCCATTTCTCCGGCGGGGTTCCACCACAGAGTGTGTGAGTTAT 2873
Oy 329 ----- 329
Db 2874 GGTGTGACTGTGTGGAGCTGATGACTTTTGGGCCAAACCTTACGATGATCCAGCC 2933
Oy 329 ----- 329
Db 2934 CCGGAGATCCCTGACCTCTGGAAAAAGGGGAGCGGCTGCCAGCCCCCATCTGTACC 2993
Oy 329 ----- 329
Db 2994 ATTGATGCTACATGATCATGTGCAAAATGTTGGATGATTGACTGTGAATGTGGCAAGA 3053
Oy 329 ----- 329
Db 3054 TTCCGGAGTTGTGTGTGAATTTCCCGCATGCCAGGAGACCCAGCGCTTGTGTGTC 3113
Oy 329 ----- 329
Db 3114 ATCCAGAAATGAGACTTGGGCCAGACAGTCCCTTGGACAGCACTGTACCGCTCACTG 3173
Oy 329 ----- 329
Db 3174 CTGAGAGCATGACATGGGGGACCTGTGTGATGCTGAGAGATATCTGTACCCAGCAG 3233
Oy 330 ----- 340
Db 3234 GGCTTCTTGTCCAGACCTGCCCCGGCGCTGGGGGCAATGTCTCACACAGGACGCG 3293
Oy 341 SerSerSerThrArgSerGlyGlyLysPleuThrLeuGlyLeuGluProSerGlu 360
Db 3294 AGCTCATCTACAGAGAGTGGGGGTGGGACCTGACACTAGGGGCTGTGAGACCTCTGAAAG 3353

Db 1225 GCCAATATCCAGAGCTTGTGGCTGCAAGAAGATCTTTGGAGCCTGGCATTTCTGCCG 1284
Oy 329 ----- 329
Db 1285 GAGAGCTTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCAGAGCACTGCCAA 1344
Oy 329 ----- 329
Db 1345 GTGTTGAGACTCTGGAGAGATCAGAGTTACTATACATCTCAGCATGGCCGGAGAGC 1404
Oy 329 ----- 329
Db 1405 CTGCTGACCTCAGCGCTCTCCAGAACCTGCMAATATCCGGGAGCAATTCGCACAA 1464
Oy 329 ----- 329
Db 1465 GCGCGCTACTGCTGACCCCTGCAGAGGCTGGGCATCAGCTGGCTGGGCTCGCTCAGCTG 1524
Oy 329 ----- 329
Db 1525 AGGGAATGGGCACTGAGACTGGCCCTCATCCACATTAACCCACCTCTGCTTGGTCAG 1584
Oy 329 ----- 329
Db 1585 ACGTGCCCTGGGACCACTCTTTGGAACCCGACCAAGCTCTGCTCCACACTGCCAAC 1644
Oy 329 ----- 329
Db 1645 CCGCCAGAGAGAGTGTGTGGGCGAGGCGCTGGCCCTGCCACACCTGTGCGCCGAGAG 1704
Oy 329 ----- 329
Db 1705 CACTGCTGGGGTCCAGGGCCCAACCAGTGTGTCAACTGACAGCAGTTCTTGGGGCCAG 1764
Oy 329 ----- 329
Db 1765 GAGTGCCTGAGAGATGCCGAGTACTGCAGAGGGCTCCCGAGGAGTATGTGAATCCAGG 1824
Oy 329 ----- 329
Db 1825 CACTGTTTGCCTGCCACCCTGAGTGTGAGCCCAAGATGGCTGAGTGACTGTTTGA 1884
Oy 329 ----- 329
Db 1885 CCGAGGCTGACCACTGTGTGCTGTGCCCACTATAAGAACCTTCCCTTGTGCGTGGCC 1944
Oy 329 ----- 329
Db 1945 CGCTGCCCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATGTGGAAGTTCCAGAT 2004
Oy 329 ----- 329
Db 2005 GAGAGGGCGCATGCGACCTTGCCCATCACTGACCCCACTCCTGTGTGAGACCTGGAT 2064
Oy 329 ----- 329
Db 2065 GACAGAGGCTGCCCGCGAGCAGAGACCACTCTGAGCTCCATGCTCTGTGCGGTG 2124
Oy 329 ----- 329
Db 2125 GTTGGCATTTCTGCTGTGCTGTGTTGGGGGTGTCTTTGGGATCCTCATCAAGCAGAG 2184
Oy 329 ----- 329
Db 2185 CAGCAGAAAGATCCGGAAGTACAGATGGGGAGACTGTGCAGAAACGAGAGCTGTGGAG 2244
Oy 329 ----- 329
Db 2245 CCGCTGACACCTAAGGGAGAGCATGGCCAAACGAGGGCAGATCGGATCTCTGAAGAAGAGC 2304
Oy 329 ----- 329

Db 2305 GAGCTGAGAGAGTGAAGTGTGATGTGCGCTTTTGGCAGACTACAGGGCATC 2364
Oy 329 ----- 329
Db 2365 TGGATCCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAGTGTGAGGGAAC 2424
Oy 329 ----- 329
Db 2425 ACATCCCCAAGCCAAAGAAATCTTAGCAGAGCATACGTATGGCTGTGTGGGC 2484
Oy 329 ----- 329
Db 2485 TCCCATATGTCTCCGCTTCTGGGCATCTGCTGACATCCAGCTGACCTGTGACA 2544
Oy 329 ----- 329
Db 2545 CAGCTTATGCCCTATGGCTGCCCTTAGACCATGTCCGGAAACCGGAGCGCTGGGC 2604
Oy 329 ----- 329
Db 2605 TCCAGACCTGCTGACTGTGTATGCAGATTGCCAAGGGAGTACCTGAGAGAT 2664
Oy 329 ----- 329
Db 2665 GTGCGCTCGTACACAGGAGCTTGGCGCTCGAAGCTGTGTCAAGATCCCAACCAT 2724
Oy 329 ----- 329
Db 2725 GTCAAAATTAACAGACTTGGGGCTGGCTGCTGCTGACATTGACAGACAGATACAT 2784
Oy 329 ----- 329
Db 2785 GCAGATGGGGGCAAGTGCCTCATCAAGTGAATGGCGCTGAGTCCATTCTCCGCGGGG 2844
Oy 329 ----- 329
Db 2845 TTACCCACCAAGATGATGTGAGATTATGGTGTGACTGTGTGGAGCTGATGACTTT 2904
Oy 329 ----- 329
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTCTGGAAGAGGG 2964
Oy 329 ----- 329
Db 2965 GAGCGGCTGCCCCAGCCCCCATCTGCACCATTTGATGTCTACATGATGTCATAATGT 3024
Oy 329 ----- 329
Db 3025 TEGATGATTACTGTGAATGTGCGCAAGATTCCGGAGTTGTGTGAATTCGCCGC 3084
Oy 329 ----- 329
Db 3085 ATGCCAAGGACCCCGAGCGCTTGTGTGTCATCCAGAAATGAGACTTGGGCCAGCCAGT 3144
Oy 329 ----- 329
Db 3145 CCGTTGACAGACCTTACCGCTCACTGCTGAGAGACGATGACATGGGGAGCTGGTG 3204
Oy 330 ----- 330
Db 3205 GATGCTGAGAGATATCTGTATCCAGAGGCTTCTTGTCCAGACCTGCCCGGGC 3264
Oy 331 AAGTGTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3324
Db 3265 GCTGGGGGATGTGCTCAGCAGGAGCGGACCTTATCTACAGAGAGTGGCGTGGGAGC 3384
Oy 331 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProAlaGlySerProLeuAlaProSer 370
Db 3325 CTGACACTAGGGCTGAGCCCTCTGAAGAGAGGCCCCAGAGTCCACTGACACCTCC 384
Oy 371 GluGlyAlaGlySerAspAlaPheAspGlyLeuGlyMetGlyAlaAlaGlySerLeu 390
Db 385 GAAAGGGCTGGCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3444

Db 2637 TACCTTTTACGAGCTGGGATCTGCACCTCGTCTGCCCTGCACAAACAGAGTG 2696
QY 329 ----- 329
Db 2697 ACAGCAGAGATGAACACAGCGGTGTAGAAAGTCAGCAAGCCCTGTGCCGAGTGTGC 2756
QY 329 ----- 329
Db 2757 TATGCTGTGGGATGAGACACTTGGAGAGGTGAGGGCAGTTACGAGTCCATATCCAG 2816
QY 329 ----- 329
Db 2817 GAGTTTGTGCTGCAGAAAGATCTTTGGAGCCTGGCATTTCTGCCGAGAGCTTTGAT 2876
QY 329 ----- 329
Db 2877 GGGGACCCAGCCTCCAACTGCCCCGCTCAGCCAGAGCAGCTCCAAAGTTTGAAGACT 2936
QY 329 ----- 329
Db 2937 CTGGAAGAGATCAGAGTTACCTATATCATCTCAGCATGGCCGGACAGCCTGGCTGACTC 2996
QY 329 ----- 329
Db 2997 AGCGTCTTCCAGAACCTGCAGTAATCCGGGAGAGATTTCTGCACAAATGGCGCTACTCG 3056
QY 329 ----- 329
Db 3057 CTGACCCCTGCAGAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCACTAGAGGAACCTGGGC 3116
QY 329 ----- 329
Db 3117 AGTGAAGCTGGCCCTCATCCACCATAAACACCACCTCTGCTTGCTGCACAGGTCCTGG 3176
QY 329 ----- 329
Db 3177 GACCAAGCTCTTTCGGAACCCGACCAAGCTCTGCTCCACACTGCCAACCAGGCGAGAGAC 3236
QY 329 ----- 329
Db 3237 GAGTGTGTGGGGAGGGCTGGCTGCTCCACAGCTGTGCGCCGAGAGGCACTGCTGGGT 3296
QY 329 ----- 329
Db 3297 CCAGGGCCCAACCAAGTGTCAACTGCAGCAGTTCTTCGGGGCCAGAGTCCCTGGAG 3356
QY 329 ----- 329
Db 3357 GAATGCCAGTACTGCAGGGGCTCCCCAGGAGTATGTAAATGCCAGGCACTGTTTGGCG 3416
QY 329 ----- 329
Db 3417 TGCCACCTGAGTGTACGCCCCAGAAATGGCTCAGTACCTGTTTGGACCAGGAGCTGAC 3476
QY 329 ----- 329
Db 3477 CAGTGTGTGGCCTGTGCCACTATAGACCCCTCCCTTCGCGTGGCCGCTGCCACAG 3536
QY 329 ----- 329
Db 3537 GGTGTGAAACCTGACCTCTCTACATGCCACTGTGAAGTTCCAGATGAGAGAGGCGCA 3596
QY 329 ----- 329
Db 3597 TGCCAGCCTTGCCCATCAACTGCAACCACTCTGTGTGAGACTGATCAGAAAGGCTGC 3656
QY 329 ----- 329
Db 3657 CCGCGGAGAGAGAGCAGACCCCTGTGAGCTCAGTCTCTGCGGTGTGACATTCTG 3716
QY 329 ----- 329
Db 3717 CTGCTGTGTGCTTGGGGTGTCTTTTGGAGTCTCATCAAGCAGAGGAGCAAGAGATC 3776

QY 329 ----- 329
Db 3777 CCGAACTACAGATGCGGAGACTGCTGCAGGAAACGAGCTGTGAGCCGCTGACACT 3836
QY 329 ----- 329
Db 3837 AGCGAGGATGCCCAACAGGCGCAGATCGGATCTGAAAGAGACGGAGCTGAGAAAG 3896
QY 329 ----- 329
Db 3897 GTGAAGTGTCTTGATCTGGCGCTTTTGGCAGAGTCTACAAAGGCACTGATCCCTGAT 3956
QY 329 ----- 329
Db 3957 GGGGAGATGTGAAAATTCAGTGGCCATCAAAAGTTTGAAGGAAAACACATCCCCCAA 4016
QY 329 ----- 329
Db 4017 GCCAACAAGAAATCTTAGACGAGCATACGTATGATGGTGTGTGGCTCCCATATATGC 4076
QY 329 ----- 329
Db 4077 TCCCGCTTCTGGGATCTGCTGACATCCAGGTGCAGCTGTGACACAGCTATGCC 4136
QY 329 ----- 329
Db 4137 TATGGCTGCTCTTAGACATGTCCGGGAAAAACCGGACGCTGGGCTCCAGACACTG 4196
QY 329 ----- 329
Db 4197 CTGAAGTGTATGACAGATTCGAAGGGATGAGTACCTGAGAGATGTGGGCTGTA 4256
QY 329 ----- 329
Db 4257 CACAGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACCAATGCAAAATTACA 4316
QY 329 ----- 329
Db 4317 GACTTCGGGCTGGCTCGCTGCTGACATTGACGAGACAGATACCATCAGATGGGGGC 4376
QY 329 ----- 329
Db 4377 AAGTTCACATGAAGTATGATGCGCTGAGTCCATTCTCCGCGGCTTACACCACAG 4436
QY 329 ----- 329
Db 4437 AGTATGTGTGAGTTATGGTGTGACTGTGTGGAGCTGATGACTTTGGGCGCAAACT 4496
QY 329 ----- 329
Db 4497 TACGATGGATCCAGCCCGGAGATCCCTGACCTGTGAAAAGGGGAGGGGCTGCC 4556
QY 329 ----- 329
Db 4557 CAGCCCCCATCTGCACATGATGTCTACATGATCATGTGAAATGTTGATGATTGAC 4616
QY 329 ----- 329
Db 4617 TCTGAATGTGGCCAAAGATCCGGGAGTTGTGTCTGAATTCTCCGCAATGCCAGGAC 4676
QY 329 ----- 329
Db 4677 CCCCAGCCTTGTGTGATCCAGAAATGAGACTTGGGCCAAGCAATCCCTTGACAGC 4736
QY 329 ----- 329
Db 4737 ACCTTACCGCTCACTGCTGAGAGAGATGACATGAGGGGACTGTGTAAGCTGAGAG 4796
QY 330 ----- 334
Db 4797 TATCTGTACCCAGCAGGCTTCTTCTGTCCAGACCTGCCCGGGCGCTGGGGCATG 4856

```

Oy 335 ValHSHARGHASrSerSerSerThrArgSerGlyGlyAspLeuThrLeuGly 354
Db 4857 GTCCACCAAGGACCGACGATCTATCTACAGAGTGGCGGGGAGCTGACCTAGG 4916
Oy 355 LeuGIuProSerGIuGIuAlaProArSerProLeuAlaProSerGIuGIuAlaGly 374
Db 4917 CTGGAGCCCTCTGAGAGAGAGGCCCCAGCTCTCCACTGCGACCTCCGAGGGCTGCG 4976
Oy 375 SerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuPro 394
Db 4977 TCCGATGATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5036
Oy 395 ThrHisAspProSerProLeuGlnArgTyrSerGIuAspProThrValProLeuProSer 414
Db 5037 ACACATGACCCCAAGCCCTCTACAGCGGTACAGTGAAGGACCCACAGTACCTGCTGCTT 5096
Oy 415 GluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlnTyrValAsnGln 434
Db 5097 GAGACTGATGGCTACGTTGCCCCCTGACCTGACGACCCCAAGCTGAATATGTGAACAG 5156
Oy 435 ProAspValArgProGlnProProSerProArGIuGIuProLeuProAlaAlaArgPro 454
Db 5157 CCAGATGTTCCGGCCCAAGCCCTCTGCGCCGAGAGGGCCCTGCTGCTGCTGCTGCTGCT 5216
Oy 455 AlaGIuAlaThrLeuGIuAlaArgAlaLeuThrLeuSerProGIuGIuAlaValAlaGly 474
Db 5217 GCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5276
Oy 475 AspValPheAlaPheGlyGlyAlaValGIuAsnProGIuTyrLeuThrProGlnGly 494
Db 5277 GACGTTTGGCTTGGGGGTGCGGTGAGAACCCCGAGTACTTGTACACCCCAAGGAGGA 5336
Oy 495 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 514
Db 5337 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5396
Oy 515 TrpAspGlnAspProProGIuAlaArgGIuAlaProProSerThrPheGlyGlyThrProThr 534
Db 5397 TGGGACCAAGGACCCCAAGAGCGGGGCTCCACCAAGCTTCAAGAGGACCTTACG 5456
Oy 535 AlaGIuAsnProGIuTyrLeuGIuLeuAspValProAlaAlaAlaProAlaArgSerPro 554
Db 5457 GCAGAGAACCCAGAGTACTGCGGTCTGAGCTGCCA-----GTGTGAACAGAGAAGGCCA 5510
Oy 555 SerProSerThrGlnProTyrGluHisValAsnAlaAlaGlnGIuAlaArgArgLeuLeu 574
Db 5511 AGTCCGCGAG---AAGCCCTGATGTCTCTCAGG-----GAGCAGAGGAGAGCGGCTCTG 5561
Oy 575 AsnLeuSerArgAspThr 580
Db 5562 AGCTATTCAGAGAGTACT 5579

```

```

RESULT 9
US-09-441-411-5
: Sequence 5, Application US/09441411
: Publication No. US2003008342A1
: GENERAL INFORMATION:

```

```

: APPLICANT: Scholler, Nathalie B.
: APPLICANT: Disis, Mary L.
: APPLICANT: Helstrom, Ingegerd
: APPLICANT: Helstrom, Karl Erik
: TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
: FILE REFERENCE: 730033.409
: CURRENT APPLICATION NUMBER: US/09/441.411
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 5
: LENGTH: 4473
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-441-411-5

```

```

Alignment Scores:
Pred. No.: 3,44e-200 Length: 4473
Score: 2418.00 Matches: 523
Percent Similarity: 41.06% Conservative: 5
Best Local Similarity: 40.67% Mismatches: 22
Query Match: 64.67% Indels: 736
DB: 9 Gaps: 4

```

US-09-821-883-2 (1-690) x US-09-441-411-5 (1-4473)

```

Oy 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 118 CGCGCCCTCCCAAGCCGGGTCCAGCCGGAGCCATGCGCGGAGCCGAGTACAGCACCATG 177
Oy 18 PheLeuPheLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 37
Db 178 GAGCTGGGGGGCTGTGCGCGTGG-----GGGCTCTCTGCGCGCTTGGCCCCC 228
Oy 38 GlyAlaAlaSerThrGlnValAlaCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 229 GGAGCCGCGACACCCCAAGTGTGACCCGACAGACATGAAGCTGCGGCTCCCTGCGCAGT 288
Oy 58 ProGIuThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnGly 77
Db 289 CCCGAGACCCCAAGCTGACATGCTCCGACCTTACCAAGGCTGCGAGGTGCGAGGA 348
Oy 78 AsnLeuGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 349 AACCTGGAACCTACCTTACCTGCGCCCAACCAATGCCAGCTGCTCTCTCTCTCTCTCTCTCT 408
Oy 98 GluValGlnGlyTyrValLeuLeuLeuHisAsnGlnValArgGlnValProLeuGlnArg 117
Db 409 GAGGTGACGGGCTACGCTGCTATGCTGCTACCAACCAAGGAGGAGGCTCCACTGCGAGG 468
Oy 118 LeuArgIleValAlaGlyGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAsp 137
Db 469 CTGGGAGATGTCGACGAGGACGACCCAGCTCTTGAAGACAACTTGCCCTGCGCTGCTGCTGCT 528
Oy 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGIuGIuLeuArg 157
Db 529 AATGAGACCCCGTGAACAAATACCAACCCCTGTACAGAGGCGCTCCCGAGAGCGCTGCGG 588
Oy 158 GluLeuGlnLeuArgSerLeuThrGlnLeuLeuGlyGlyValLeuLeuIleGlnArgAsn 177
Db 589 GAGCTGACGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 648
Oy 178 ProGlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln 197
Db 649 CCCAGCTCTGCTACACAGGACAGCATTTTGTGAAGGACATCTTCCACAAACACACAG 708
Oy 198 LeuAlaLeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerPromet 217
Db 709 CTGCTCTCAACATGATAGACACCAACCGCTGCGGCTGCTCCACCCCTGCTCTCCGATG 768
Oy 218 CysLysGlySerArgCysTyrGlyGlyLeuSerGlnAspCysGlnSerLeuThrArgThr 237
Db 769 TGTAAAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
Oy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGln 257
Db 829 GTCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 889 CAGTGTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
Oy 278 AsnHisSerGlyIleCysGlyLeuHisCysProAlaLeuValIleTyrAspThrAspThr 297
Db 949 AACCAAGTGGCATCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008
Oy 298 PheGluSerMetProAsnProGIuGIuArgTyrThrPheGlyAlaSerCysValThrAla 317
Db 1009 TTTGAGTCCATGCCCATTCGAGGCGCGGTATATCATTCGCGGCGCTGCTGCTGCTGCTGCT 1068

```

QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
Db 1069 TETCCCTACAACTACTCTTCTACGGACGTGGATCTCTGCACCCCTGCTGCCCCCTGCAC 1128
QY 329 ----- 329
Db 1129 AACCAAGAGTGCACAGAGAGTGAACACAGCGGTGTGAGAAATGTCAGCAAGCCCTGT 1188
QY 329 ----- 329
Db 1189 GCCCAGTGTGCTATGTGCTGGCATGGACACTTGCAGAGGTGAGGGCAGTTACCACT 1248
QY 329 ----- 329
Db 1249 GCCAATATCCAGAGATTCTGCTGCTGCAGAAAGATCTTTGGAGCCTGGCATTTCTGCCG 1308
QY 329 ----- 329
Db 1309 GAGAGCTTGTATGGGAGCCACCTCCACACTGCCCGCTCCAGCCAGAGAGCTCCAA 1368
QY 329 ----- 329
Db 1369 GTGTTGAGACTCTGGAAGAGATCACAGGTTACTATACATCTCAGCATGGCCGGACAGC 1428
QY 329 ----- 329
Db 1429 CTGCCTGACCTGAGGCTTCTCCAGAACTGCAGAAATACCGGGAGCAATTTCTGCACAA 1488
QY 329 ----- 329
Db 1489 GCGGCTTACTGCTGACCCCTGCAGAGGGCTGGGCAATCAGCTGCTGGGCTGCTCAGTG 1548
QY 329 ----- 329
Db 1549 AAGGAATGGGAGTGGAGTGGCCCTCATCCACCAATACACCCACCTCTGCTTCGTGCAC 1608
QY 329 ----- 329
Db 1609 ACGGTGCCCTGGGACCAAGCTTTTGGGAACCCGCAACCAAGCTCTGCTCAGACTGCCAAC 1668
QY 329 ----- 329
Db 1669 CGGCCAGAGAGAGTGTGTGGCGAGGGCTGGCTGCCACCAAGCTGTGCCGCCGAGGG 1728
QY 329 ----- 329
Db 1729 CACTGCTGGGTCCAGAGGCCACCAAGTGTCACTGACGCCAGTTCTTGGGGCCAG 1788
QY 329 ----- 329
Db 1789 GAGTGCCTGGAGGAATGCCGAGTACTGTCAGAGGGCTCCCAAGGAGTATGTGAATGCCAG 1848
QY 329 ----- 329
Db 1849 CACTGTTTCGCTGCCACCCCTGAGTGTCAAGCCCAAGATGGCTCAGTGAATGTTTGA 1908
QY 329 ----- 329
Db 1909 CCGAGGCTGACCAAGTGTGTGGCTGTGCCCAATATAGAGACCTCCCTTCTGCGTGGCC 1968
QY 329 ----- 329
Db 1969 CCCTGCCCAAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGAT 2028
QY 329 ----- 329
Db 2029 GAGAGGGGCGATGCCAGCCTTGCCCATCACTGCACCCACTCTGTGTGACCTGAT 2088
QY 329 ----- 329
Db 2089 GACAAAGGCTGCCCGCCGAGAGAGAGCCACCTCTGAGCTCATCATCTCTGCGGTG 2148

QY 329 ----- 329
Db 2149 GTTGGCATTTCTGCTGTGTGCTTTGGGGGTGTCTTTGGATCCTATCAAGGAGCG 2208
QY 329 ----- 329
Db 2209 CAGCAGAGATCCGGAGTACAGATGCGAGACTGTGACAGAAACGAGCTGTGTAG 2268
QY 329 ----- 329
Db 2269 CCGCTGACCTTAGCGGAGCGATGCCCAACAGCGCGAGATCGGATCCTGAAGAGACG 2328
QY 329 ----- 329
Db 2329 GAGCTGAGGAAGTGAAGTGTGTGATCTGGCGCTTTTGGCAGTGTACAAAGGCAATC 2388
QY 329 ----- 329
Db 2389 TGGATCCCTGATGGGAGATGTGAATAATCCAGTGGCCATCAAGTGTGAGGAAAC 2448
QY 329 ----- 329
Db 2449 ACATCCCCCAAGCCACAAAGAAATCTTAGACGAAGCATACGTGATGGCTGTGGGC 2508
QY 329 ----- 329
Db 2509 TCCCATATGTCTCCCGCTTGGGCAATCTGCTGACATCCAGGCTGCAGCTGGTGACA 2568
QY 329 ----- 329
Db 2569 CAGTTATCCCTATGGCTGCTTAGACATGTCCGGGAAACCGCGAGCGCTGGGC 2628
QY 329 ----- 329
Db 2629 TCCAGAGACTGCTGAATGTGTATGATGAGATTGCCAAGGGATGACTACTGGAGAT 2688
QY 329 ----- 329
Db 2689 GTGGGCTGTACACAGGAGCTTGGCGCTCGGAAACGTGCTGATCAAGATCCCAACAT 2748
QY 329 ----- 329
Db 2749 GTCAAAATTCAGACTTCGGGCTGCTCGCTGTGACATTGACAGAGACAGTACCAT 2808
QY 329 ----- 329
Db 2809 GCAGATGGGGCAAGTGGCCATCAAGTGAATGGCGCTGAGTCAATTTCCGCCGCCGG 2868
QY 329 ----- 329
Db 2869 TTCACCCACAGAGTATGTGTGAGTTATGTTGATGTGTGGAGCTGATGACTTTT 2928
QY 329 ----- 329
Db 2929 GGGGCCAAACTTACGATGGATCCAGCCCGGAGATCCCTGACTGTGAAAAAGGG 2988
QY 329 ----- 329
Db 2989 GAGCGGCTCCCGAGCCCGCCCATCTGCACCAATTGATGTCTACATGATCATGTCAAAATGT 3048
QY 329 ----- 329
Db 3049 TGGATGATTGACTGTGAATGTGGGCCAAGATTCCGGGATTTGTTGTCTGAATTTCCCGC 3108
QY 329 ----- 329
Db 3109 ATGGCAGGGAGCCCGAGCGCTTGTGTATCAGAAATGAGGACTTGGGCCAGCAGT 3168
QY 329 ----- 329
Db 3169 CCCTTGGACGACCTTCTACCGCTCAGCTGTGAGAGACGATGACATGGGGAGCTGTG 3228
QY 330 -----Gly 330

QY 329 ----- 329
Db 946 CCCCTGCACAACAAGAGTGACAGAGATGGAAACACGCGTGTGAGAAAGTGACG 1005
QY 329 ----- 329
Db 1006 AAGCCCTGTGCCAGAGTGTCTATGTCTGGGATGAGACCTTGCAGAGGTGAGGGCA 1065
QY 329 ----- 329
Db 1066 GTTACCAATGCCAATATCCAGAGAGTTGTGTGCTGCAGAAAGATCTTTGGAGCCTGCA 1125
QY 329 ----- 329
Db 1126 TTTCGCCGAGAGAGCTTTGATGGGAGCCACCTCCAACTGCCCCGCTCCAGCCAG 1185
QY 329 ----- 329
Db 1186 CAGCTCCAAAGTTTGAAGACTGTGGAGAGATCACAGGTTACTATACATCTCAGCATGG 1245
QY 329 ----- 329
Db 1246 CCGGACAGCTTGCCTGACCTCAGCGTCTTCCAGAACTGCAGATATCCGGGAGCAATT 1305
QY 329 ----- 329
Db 1306 CTGCACATGGCGCTACTGCTGACCCCTGCAAGGCTGGGCAATCAGCTGGGCGTGG 1365
QY 329 ----- 329
Db 1366 CGCTACTGAGGAACTGGGAGTGGAGTGGAGTGGCCCTCATCACATACACCACCTCTGC 1425
QY 329 ----- 329
Db 1426 TTCTGCACACAGTGCCCTGGGACACGCTTTTGGAAACCGACCAAGCTCTGCTCAC 1485
QY 329 ----- 329
Db 1486 ACTGCACAACGGCCAGAGAGAGAGTGTGTGGGCGAGGCGCTGGCCTGCACAGCTGTC 1545
QY 329 ----- 329
Db 1546 GCCGAGGCACTGTGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCAGTTCTT 1605
QY 329 ----- 329
Db 1606 CGGGCCAGAGTGTGGAGAAATCCGAGTACTGCAGGGGCTCCAGAGGATATGTG 1665
QY 329 ----- 329
Db 1666 AATGCACAGCACTTTTGCCTGCCAACCCTGAGTGTCAAGCCCAAGATGGCTCAGTGC 1725
QY 329 ----- 329
Db 1726 TGTTTGGACCGAGGCTGACCAAGTGTGTGCTGTGCCACTATAGAGACCTTCTTC 1785
QY 329 ----- 329
Db 1786 TGGGTGGCCGCTGCCCAAGGCTGTGAACCTGACCTCTCTACATGCCATCTGGAAG 1845
QY 329 ----- 329
Db 1846 TTTCAGATGAGAGAGGCGCATGCGAGCTTGGCCCATCACTGACCCACACTCTGTGTG 1905
QY 329 ----- 329
Db 1906 GACCTGATGACAAAGGCTGCCCGCGAGAGAGACCAAGCCCTGTGACGTCAATCGTC 1965
QY 329 ----- 329
Db 1966 TCTGGGCTGTGGCATTTCTGTGTGCTCTTGGGGGTGTCTTTGGGATCTCTATC 2025
QY 329 ----- 329

Db 2026 AAGCAGCGGACAGAAAGATCCGGAAGTACAGATGCGGAGACTGTCCAGAAAAAGAG 2085
QY 329 ----- 329
Db 2086 CTGTGAGCCGCTGACACCTAGCGGAGCGATGCCAACAGCGCGAGATCGATCTGTG 2145
QY 329 ----- 329
Db 2146 AAAGAGAGAGCTGAGGAAGGTGAAGTGTGTGATCTGGCGCTTTTGGCAGACTCAC 2205
QY 329 ----- 329
Db 2206 AAGGCACTGTGATCCCTGATGGGAGAAATGAAAAATTCAGTGGCCATCAAAATGTGTG 2265
QY 329 ----- 329
Db 2286 AAGGAAACACATCCCGCAAGCAAAAGAAATCTTAGAGAAAGCATACGTATGGCT 2325
QY 329 ----- 329
Db 2326 GGTGTGGCTCCCATATGTCTCCCGCTCTGGGCAATCTGCTGACATCCAGGTCAG 2385
QY 329 ----- 329
Db 2386 CTGTTGACACAGCTTATGCTATGCTGCTCTTAGACATGTCCGGAAAAACCGCGGA 2445
QY 329 ----- 329
Db 2446 CGCTGGGCTCCAGAGACCTGCTGAACGTGTATGAGAAATGCCAAGGAGATGACTAC 2505
QY 329 ----- 329
Db 2506 CTGAGAGATGTGGCTGTACACAGGACCTTGGCCGCTCCGAAAGCTCTGTCAAGAGT 2565
QY 329 ----- 329
Db 2566 CCACAACATGTCAAAATATACAGACTTGGGCTGGCTCGCTGTGACATTGACGAGACA 2625
QY 329 ----- 329
Db 2626 GAGTACATGAGATGGGGGAGAGGTGCCATCAAGTGAATGGCGTGAATCTCTC 2685
QY 329 ----- 329
Db 2686 CGCGGCGCTTACCCACACAGATGATGTGGAATATGTGTACTGTGTGGAGCTG 2745
QY 329 ----- 329
Db 2746 ATGACTTTTGGGGCCAAACTTACGATGGATCCAGCCGAGAGATCCCTGACCTGCTG 2805
QY 329 ----- 329
Db 2806 GAAAAAGGGAGCGGCTCCCGAGCCCGCATCTGCACCAATGATGTACATGATCATG 2865
QY 329 ----- 329
Db 2866 GTCAAAATGTGGATGATTTGACTCTGAATGTGCGGCCAAGATTCGGGAGTTGTGTCTGA 2925
QY 329 ----- 329
Db 2926 TTCTCCGATGAGCGAGGAGCCCGAGCGCTTTGTGTATTCAGAAATGAGAGACTTGGGC 2985
QY 329 ----- 329
Db 2986 CCAGCAAGTCCCTTGACAGACACTTCTACCGCTCATCTGTGAGAGAGATGACATGGGG 3045
QY 329 ----- 329
Db 3046 GACCTGTGATGTGAGGAGATATCTGTACCCAGAGGCTTCTGTGTCCAGACCT 3105
QY 330 ----- 347
-----GlyAlaGlyGlyMetValHisHisArgHisArgSerSerThrArgSerGly
|||||

```

Db 3106 GCCCGGGGCGCTGGGGGCAATGGTCCACACAGGACACCGGACGACTCATTTACACAGAGGTGGC 3105
Oy 348 GtlytlyspLeuthrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeu 367
Db 3166 GGTGGGGGACCTGACACTAGAGGGCTGGAGCCCTCTGAAAGAGAGAGCCCGACAGTCTCCACTG 3225
Oy 368 AlaProSerGluGlyAlaGlySerAspAlaPheAspGlyAspLeuGlyMetGlyAlaLa 387
Db 3226 GCACCTCTCCGAAGGGGCTGGCTCCGATGTAATTTGATGATGGACTGGAGTAATGGGGGACCC 3285
Oy 388 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAsp 407
Db 3286 AAGGGGCTGCMAAGCCCTCCACACATGACCCAGCCCTCTACAGGGTACAGTATGAGAC 3345
Oy 408 ProthrValProLeuProSerGluThrAspGlyTyrValAlaProLeuthrCysSerPro 427
Db 3346 CCCACAGTACCCCTGGCCCTCTGAGACTGATGGCTACGTTGGCCCTGACCTGCAGCCCC 3405
Oy 428 GlnProGluTyrValaIngInProAspValArgProGlnProProSerProArgGluGly 447
Db 3406 CAGCTCAATATGTGAACCGACCGAGATGTTGGCCCGCCCGCCCTTCGCGCGGAAGGGC 3465
Oy 448 ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerPro 467
Db 3466 CCTGCGCTGCTGGCCGACCTGGTGGCGCACTGTGAAGAGGCCAAGACTCTCTCCCA 3525
Oy 468 GlytlysaNglyValValLysAspValPheAlaPheGlyGlyAlaValGluAspProGlu 487
Db 3526 GGGAGAGATGGGGTGGCTCAAGACGTTTTTGGCTTTGGGGGGCGCTGGAGAACCCCGAG 3585
Oy 488 TyrLeuthrProGlnGlyGlyAlaAlaProGlnProHisArgProProAlaPheSerPro 507
Db 3586 TACTTGACACCCCGAGGAGAGAGCTGCCCTCAGCCCCACCTCTCTGCTTACAGCCA 3645
Oy 508 AlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSer 527
Db 3646 GCCTTCGCAACCTCATTTACTGCGGAGACAGAGACCCAGACCGGGGGGCTCCACACAGC 3705
Oy 528 ThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValPro 546
Db 3706 ACCTTCAAAAGGACACTTACGGCAGAGAACCCAGAGTACTGTGGTGTGACGTGCCA 3762

RESULT 11
US-09-811-115-2
; Sequence 2, Application US/09811115
; Patent No. US2002003736A1
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT. 03A4
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3768
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-811-115-2

```

Alignment Scores:

Pred. No.:	1,25e-199	Length:	3766
Score:	2410.50	Matches:	510
Percent Similarity:	41.16%	Conservative:	0
Best Local Similarity:	41.16%	Mismatches:	2
Query Match:	64.47%	Indels:	727
DB:	10	Gaps:	1

US-09-821-883-2 (1-690) x US-09-811-115-2 (1-3768)

QY	35	LeuAlaArgGlyAlaAlaSerThrGlnValCysThrGlyThrAspMetCysLeuArgLeu	54
Db	46	TTGGCCCCCGGAGACCGGGAGACCCCAAGTGTGACCGGACACATGAAGCTGGCGCTC	105
QY	55	ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyroGlnCysGlnVal	74
Db	106	CTGCGACATCCCGAGACCACCTGGACATGCTCCGACACTGCATACAGAGGCTGCAGGTTG	165
QY	75	ValGlnGlyAsnLeuGluLeuThrTyroLeuProThrAsnAlaSerLeuSerPheLeuGln	94
Db	166	GTGAGGGGAAACCTGGAACTCACCATTCTGCCACCAATGCCAGGCTGTCTCTCCGACG	225
QY	95	AspIleGlnGluValGlnGlyTyroValLeuIleAlaHisAsnGlnValArgGlnValPro	114
Db	226	GATATCCAGAGGTGTCAGAGGCTACGTGCTCATGCTGCATACACCAAGTGAAGGATGCCCA	285
QY	115	LeuGlnArgLeuAlaArgIleValAlaArgGlyThrGlnLeuPheGluAspAsnTyroAlaLeuAla	134
Db	286	CTGCAGAGCGTGGGATTTGGCAGAGGACCCACACTCTTTTGAGACAACTATCCCTGTGCC	345
QY	135	ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGly	154
Db	346	GTGCTAAGACATGGAGACCCGCTGTAAATACACCCCTGTGCAGAGGGCTCTGCCAGGA	405
QY	155	GlyLeuArgGluLeuGlnLeuAlaArgSerLeuThrGluIleLeuLysGlyGlyValLeuIle	174
Db	406	GGCCGCGGGGAGCTGCAGCTTCGAAAGCCTCACAGAAATCTTGAAAGAGGGTCTGTGATC	465
QY	175	GlnArgAspProGlnLeuAsCysTyroGlnAspThrIleLeuThrLysAspIlePheHisLys	194
Db	466	CAGCGGAACCCCGACGCTCTGTCACAGACACACATTTTGGAGAAGACATCTTGCACAG	525
QY	195	AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys	214
Db	526	AACAACCAAGCTGGCTCTCACACTGATAGACACCAACCCGCTCTGGGCTGCCACCCCTGT	585
QY	215	SerProMetCysLysGlySerArgCysTrpGlyGluSerSerGluAspGlySerLeu	234
Db	586	TTCTCGAGTGTAAAGGCTTCCCGCTGTGGGGAGAAAGTTCTGAGGATTTGTACAGAGCTG	645
QY	235	ThrArgThrValCysAlaGlyGlyCysAlaAlaArgCysAlaGlyProLeuProThrAspCys	254
Db	646	ACGGCACATGTCGTGCCGGGTGGCTGGCCCTGCACAAAGGGGCACATGCCCACTGATGTC	705
QY	255	CysHisGluGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys	274
Db	706	TGCCATGAGCGATGTGTGGCGGTGCACGGGCCCAACACACTGTGACTGCTGTGCCCTGTC	765
QY	275	LeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyroAsn	294
Db	766	CTCCACTTCAACCACTGGCACTGTGTGACTGTGCACCTGCCACCCCTGTGTACACTTAAAC	825
QY	295	ThrAspThrPheGluSerMetProAsnProGluGlyArgTyroThrPheGlyAlaSerCys	314
Db	826	ACAGACACGTTTGATGTCCATGCCCAATCCCGAAGGGCGGTAAATACATTCGGCCGACAGCT	885
QY	315	ValThrAlaCysProTyroAsnTyroLeuSerThrAspValGlySer	329
Db	886	GTGACTGCTGTCCCTACAACTACCTTTTCTACGGAGACGTGGATCCGACACCTCTGCTGTC	945
QY	329		329
Db	946	CCCCGGCACAACAAAGAGTGNACAGCAGAGAGATGGAAACACACGGGTGTGAAAGTGCAGC	1005
QY	329		329
Db	1006	AAGCCCTGTGCCGAGTGTCTATGCTTGGCATGGACACTTGCAGAGGTGAGGCA	1065
QY	329		329
Db	1066	GTTACCAAGTGCATATTCAGAGAGTTTGTGCTGGCTGCAGAGAGATCTTTGGAGAGCTGCA	1125

329 ----- 329
Db 1126 TTTTCGCGGAGAGCTTTGATGGGAGCCAGCTCCAACTGCCCCGCTCCAGCCAGAG 1185
329 ----- 329
Db 1186 CAGCTCAAGTGTGGAGACTGTGAGAGATCACAGTTACCTATCTACAGATGG 1245
329 ----- 329
Db 1246 CCGGAGACCTGCTGACCTAGCTCTTCCAGAACTGGCAAGTAATCCGGGGAGCAATT 1305
329 ----- 329
Db 1306 CTGCACATGGCGCTACTCGCTGACCTGCAAGGGCTGGCATGAGTGGGCTG 1365
329 ----- 329
Db 1366 CGCTCACTGAGGGAAGTGGGAGTGGAGCTGGCCCTCATCCACCATTAACACCACTCTGC 1425
329 ----- 329
Db 1426 TTCTGTCACACGCTGCTCCCTGGAGACAGCTTTTGGAAACCGCACCAAGCTCTGCTCAC 1485
329 ----- 329
Db 1486 ACTGCCAACCGGCGCAGAGAGAGTGTGTGGGCGAGGGCTGGCTGCCACCAAGCTGTGC 1545
329 ----- 329
Db 1546 GCCCGAGGCACTGCTGGGGTCCAGGGCCACCCAGTGTGTCACTGACAGCAGTTCTT 1605
329 ----- 329
Db 1606 CCGGCGCAGAGAGTGGCTGGAGGAATCCGAGTACTGACAGGGCTCCCCAGGAGTATGTG 1665
329 ----- 329
Db 1666 AATGCCAGGCACTTTTGGCGGTGCCACCCCTGAGTGCAGGCCCAAGATGGCTCAGTGAAC 1725
329 ----- 329
Db 1726 TGTTTTGACCGAGGCTGACAGTGTGTGGCTGTGCCACTATAAGAGACCTCCCTTC 1785
329 ----- 329
Db 1786 TGCCTGCCCCGCTGCCCGCAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAA 1845
329 ----- 329
Db 1846 TTTCCAGATGAGAGGGGCGCATGCGAGCCTTGCCTCACTGACACCACTCCTGTGTG 1905
329 ----- 329
Db 1906 GACCTGATGACAGAGGCTGCCCGCCGAGAGAGAGCCAGCCCTCTGACGTCCATCGTC 1965
329 ----- 329
Db 1966 TCTGCGGTGGTTGGCATTTCTGCTGTGCTGTGGGGGTGGTCTTTGGGATCCTCATC 2025
329 ----- 329
Db 2026 AAGGACGGCAGCAGAAAGATCCGGAAGTACAGATGGGAGACTGCTGAGAAACGGAG 2085
329 ----- 329
Db 2086 CTGCTGAGCCGCTGACACTAGCGAGCGATGCCAACAGCGCAGATGCGGATCTCTG 2145
329 ----- 329
Db 2146 AAAGAGAGGAGCTGAGGAAGGTGTTGGATCTGGCGCTTTTGGGACAGCTTAC 2205
329 ----- 329

2206 AAGGCAATCTGGATCCCTGATGGGAGATGTGAAATTCAGTGGCATCAAACTGTG 2265
329 ----- 329
2266 AAGGAAACACATCCCCCAAAACCAAGAAATCTAGACGAAGATACGTGATGCT 2325
329 ----- 329
2326 GGTGTGGGCTCCCATATGTCTCCGCCCTTCTGGGATCTGCTGACATCCAGGTCAG 2385
329 ----- 329
2386 CTGTGACACAGCTTATGGCTATGGCTGCTTACACCATGTCCGGGAAACCGCGGA 2445
329 ----- 329
2446 CGCTGGGCTCCAGAGACCTGCTGAACCTGTGTATGCAAGATTGCCAAGGGATGAGCTAC 2505
329 ----- 329
2506 CTGAGAGATGTGGGCTGTACACAGGACTTGGCCGCTCGGACGTGCTCAAGAT 2565
329 ----- 329
2566 CCCAACATGTCAAAATTACAGACTTCGGGCTGGGCTCGGCTCTGGACATTGACGAGACA 2625
329 ----- 329
2626 GAGTACATGAGATGGGGGCAAGTGCCCATCAAGTGAATGGCGCTGAGTCCATTCTC 2685
329 ----- 329
2686 CCGCGGCGGTTACCCACACAGTATGTGTGAGTATGTTGACTGTGGGAGCTG 2745
329 ----- 329
2746 ATGACTTTGGGGCCAAACCTTACATGATGGATCCAGCCGGGAGATCCCTGACTGTG 2805
329 ----- 329
2806 GAAAGGGGAGCGGCTGCCAGCCCGCCATCTGCACATTGATGTATGATGATG 2865
329 ----- 329
2866 GTCAATGTGATGATTGACTCTGATGTGGCCAAAGATTCCGGGAGTTGGTGTGAA 2925
329 ----- 329
2926 TTCTCCGATGCGCAGGAGCCCGACGCTTTGTGTATCCAGAAATGAGGATTTGGGC 2985
329 ----- 329
2986 CCAGCAGTCCCTGGAGACACCTTACCGCTCACTGAGAGACGATGACATGGG 3045
329 ----- 329
3046 GACCTGTGATGCTGAGAGATGCTGTAACCCAGCAGGCTCTTCTGTCCAGACCT 3105
330 ----- 347
3106 GCCCGGGGCGTGGGGGATGCTCCACACAGCAGCCAGCTCATCTACAGAGATGTC 3165
348 G1yG1yAspLeuThLeuG1yLeuG1yProSerG1yG1yAlaProArgSerProLeu 367
3166 GGTGGGAGCTACACTAGGGCTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3225
368 AlAProSerG1yAlaG1ySerAspValPheAspG1yAspLeuG1yMetG1yAlaAla 387
3226 GCACCTCCGAAGGGGCTGGCTCCATGTATTTGATGTGACCTGGGAATGGGGCAGCC 3285
388 LysG1yLeuG1ySerLeuProThrLysAspProSerProLeuG1yArgTyrSerG1yAsp 407
|||||

QY 329 ----- 329
Db 1006 AAGCCCTGTGCCGAGTGTGCTATGTCCTGGCATGAGACACTTGGAGAGGTGAGGGCA 1065
QY 329 ----- 329
Db 1066 GTTACAGTGCATATATCCAGAGACTTTCGTGCTGCAAGAAGATCTTTGGAGCCTGGCA 1125
QY 329 ----- 329
Db 1126 TTTCTGCCGAGAGACTTTGATGGGGACCAGACCTCCAACTGCCCCGCTCCAGCCAGAG 1185
QY 329 ----- 329
Db 1186 CAGCTCCAGTGTTTAGACTCTGGAGAGATCAGAGTTACCTATATCATCTCAGCATGG 1245
QY 329 ----- 329
Db 1246 CCGGACAGCTGTCTGACTCAGCGTCTCCAGAACCTGCAGATATCCGGGAGCATTT 1305
QY 329 ----- 329
Db 1306 CTGCACATGAGCGCTTACTGCTGACCTCGAAGGCTGGGCATCAGCTGGGCTG 1365
QY 329 ----- 329
Db 1366 CGCTCACTGAGGGAACTGGGCACTGAGACTGGCCCTCATCCACATTAACACCCACTCTGC 1425
QY 329 ----- 329
Db 1426 TTGCTGACACAGGTGCCCTGGGACACAGCTTTTCGGAACCGCACCAAGCTCTGCTCAC 1485
QY 329 ----- 329
Db 1486 ACTGCCAACCGGCCAGAGAGAGAGTGTGTGGGCGAGGGCTTGCCCTGCCACACCTGTGC 1545
QY 329 ----- 329
Db 1546 GCCCGAGGACACTGTGGGTCAGGGCCACCAAGTGTGTCACTGAGCGAGTTCTT 1605
QY 329 ----- 329
Db 1606 CCGGGCCAGAGTGCCTGGAGGAATGCCAGTACTGAGGGGCTCCAGAGGAGTATGTG 1665
QY 329 ----- 329
Db 1666 AATGCCAGGCACTGTTGCCGTGCCACCCTGAGTGTCAAGCCCCAGAAATGGCTCAGTGACC 1725
QY 329 ----- 329
Db 1726 TGTTTGGACCGAGGCTGACAGTGTGTGGCTGTGCCACATTAAGAGACCTCCCTTC 1785
QY 329 ----- 329
Db 1786 TGCCTGCCCGCTGCCAGCGGTGTGAACCTGACTCTCTACATGCCCATCTGGAAG 1845
QY 329 ----- 329
Db 1846 TTTCCAGATGAGAGGGCGCATGCGACCTTGCCCATCAACTGCACCCACTCCTGTGTG 1905
QY 329 ----- 329
Db 1906 GACCTGATGACAGAGGCTGCCCGCCGAGAGAGACGACCTCTGACGTCCATCATC 1965
QY 329 ----- 329
Db 1966 TCTGCGGTGTTGGCACTTCTGCTGTGCTGTTGGGGGTGCTTTGGGATCTCATC 2025
QY 329 ----- 329
Db 2026 AAGGAGGCGACAGAAAGATCCGGAAGTACACGATGGCGGAGACTGTGCGAGAAACGAG 2085
QY 329 ----- 329

Db 2086 CTGCTGAGCGCGCTGACACCTAGCGGAGCGATGCCAAACGAGCGCAGATCGGATCCTG 2145
QY 329 ----- 329
Db 2146 AAGAGACGAGCTGAGGAAGGTGAAGTGTGATCTGGCCTTTTGGCACAGTCTAC 2205
QY 329 ----- 329
Db 2206 AAGGCATCTGATCCCTGATGGGAGATGTGAAAATTCAGTGGCCATCAAGTGTG 2265
QY 329 ----- 329
Db 2266 AGGGAACACATCCCCAAAGCCAAAGAAATCTTAGACGAAGCATAGTGAGCT 2325
QY 329 ----- 329
Db 2326 GGTGTGGGCTCCCATATGTCCTCCGCTTCTGGGCACTGCTGACATCCACGGTCCAG 2385
QY 329 ----- 329
Db 2386 CTGTGACACAGCTTATGCCCTATGGCTGCTTACACCATGTCCGGAAACCGGGA 2445
QY 329 ----- 329
Db 2446 CGCTGGGCTCCAGACCTGCTGACTGCTGTATGAGATTGCCAAGGGAGTACCTAC 2505
QY 329 ----- 329
Db 2506 CTGAGGATGTGCGGCTGTACACAGGACTTGGCCCTCGGAACGTGTGTCAGAGT 2565
QY 329 ----- 329
Db 2566 CCCAACCATGTCAAAATTACAGACTTGGGGCTGGCTGGCTGTGAGATTGACAGACACA 2625
QY 329 ----- 329
Db 2626 GAGTACCATGAGATGGGGGCAAGTGCCCATCAAGTGAATGGCGTGAATCTCTC 2685
QY 329 ----- 329
Db 2686 CGCCGGCGTTCAACCCACAGAGTGTGTGAGTTATGTGACTGTGTGGAGCTG 2745
QY 329 ----- 329
Db 2746 ATGACTTTGGGGCCAAACCTTAGATGGATGCCAGCCGGAGATCCCTGACCTGCTG 2805
QY 329 ----- 329
Db 2806 GAAAAGGGAGCGGCTGCCACCCCCCATCTGCACATTTGATGCTACATGATCATG 2865
QY 329 ----- 329
Db 2866 GTCAATGTGGATGATGACTTGAATGTGCGCAAGATTCCGGAGTTGTTCTGAA 2925
QY 329 ----- 329
Db 2926 TTTCCCGCATGGCCAGGAGCCCCAGCGCTTTGTGGTCATCCAGAAATGAGGACTTGGGC 2985
QY 329 ----- 329
Db 2986 CCAGCAGTCCCTTGGACAGCACCTTCTACCGCTACCTGTGAGAGACATGAGG 3045
QY 329 ----- 329
Db 3046 GACCTGTGATGCTGAGAGATATCTGTACCCACAGAGGCTTCTGTCCACACCT 3105
QY 330 ----- 347
Db 3106 GCCCGGGCGCTGGGGCATGTGTCACACAGGACCGACACTCATCTACAGAGGTGGC 3165
QY 348 GlyIysAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeu 367
|||||

Db 1066 GTTACAGTCCCAATATCCAGAGTTTGTGCTGCAGAGAGATCTTTGGAGCCTGGCA 1125
QY 329 ----- 329
Db 1126 TTTCTGCCGAGAGCTTTGATGGGAGCCAGCCTCCAACTGCCCCGCTCCAGCCAGAG 1185
QY 329 ----- 329
Db 1186 CAGCTCAAGTGTGGAGACTGTGAGAGATCACAGGTTACTATACATCTCAGCATGG 1245
QY 329 ----- 329
Db 1246 CCGAGAGCCTGCTGACCTCAGCGTCTTCAGAACTGCAAGTAATCCGGGAGCAATT 1305
QY 329 ----- 329
Db 1306 CTGCACATGGGCGCTACTCCGTGACCCCTGCAGAGGCTGGGATCAGCTGGGGCTG 1365
QY 329 ----- 329
Db 1366 CGCTCACTGAGGGAACATGGGGAGTGGACTGGCCCTCATCCACCATTAACACCCACCTTGC 1425
QY 329 ----- 329
Db 1426 TTCTGTCAACAGGTGCCCTGGGACCACTTTTCGGAACCCGACCAAGCTCTGCTCCAC 1485
QY 329 ----- 329
Db 1486 ACTGCCAACCGGCGCAGAGAGAGTGTGTGGCGAGGGCTGGCTGCCACCACTGTGC 1545
QY 329 ----- 329
Db 1546 GCCCAGAGCACTGCTGGGGTCCAGGGGCCACCAAGTGTGTCACTGCAGCCAGTTCCTT 1605
QY 329 ----- 329
Db 1606 CGGGGCCAGAGTGGCTGAGAGAAATGCCAGTACTGCAGAGGGCTCCCGAGGAGATGTG 1665
QY 329 ----- 329
Db 1666 AATGCCAGGCACTGTTTGGCGTCCACCCTGAGTGCAGCCCCAGAAATGGCTCAGTGAC 1725
QY 329 ----- 329
Db 1726 TGTTTTGACCGAGAGCTGACAGTGTGTGGCTGTGCCACTATTAAGAGACCTCCCTTC 1785
QY 329 ----- 329
Db 1786 TGCGTGGCCCGTGCCTCCAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAG 1845
QY 329 ----- 329
Db 1846 TTTCCAGATGAGGAGGGCCGATGCCAGCCTTGCCCCATCACTGCACCCACTCTGTGTG 1905
QY 329 ----- 329
Db 1906 GACCTGATGACAAGGGCTGCCCGCCGAGAGAGGACGACGCTGTGACGTCAATCATC 1965
QY 329 ----- 329
Db 1966 TCTGGGTGTGTCATCTGCTGTGCTGTGCTGTGGGGGTGGTCTTTGGATCTCATC 2025
QY 329 ----- 329
Db 2026 AAGCGACGGCAGCAAGATCCGGAAGTACACGATGCGGAGACTGCTGCAGGAAACGGAG 2085
QY 329 ----- 329
Db 2086 CTGTGAGCCGCTGACACCTAGCCGAGGAGATGCCCAACGAGCGAGATGGGATCTTG 2145
QY 329 ----- 329
Db 2146 AAGAGACGAGCTGAGGAAGGTGAAGGTGCTTGATCTGGCGCTTTTGGCACAGTCTAC 2205

QY 329 ----- 329
Db 2206 AAGGCATCTGGATCCCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAACTGTG 2265
QY 329 ----- 329
Db 2266 AAGGAAACATCTCCCAAAAGCAAGAAATCTTAGAGAGACATATCGATGGCT 2325
QY 329 ----- 329
Db 2326 GGTGTGGCTCCCATATGTCTCCGCTTCTGGCATCTGCTGACATCAAGGTGCAG 2385
QY 329 ----- 329
Db 2386 CTGTGACACAGCTTATGCCCTATGGCTGTCTTAGACATGTCCGGGAAACCGCGGA 2445
QY 329 ----- 329
Db 2446 CGCTGGGCTCCAGAGACCTGCTGAATGTGTATGATGCCAAGGGAGTACATAC 2505
QY 329 ----- 329
Db 2506 CTGAGAGATGTGGGCTGTACACAGGACTTGGCCGCTCGGAAGTGTCTCAAGAGT 2565
QY 329 ----- 329
Db 2566 CCCAACATGTCAAAATTTACAGACTTCCGGGCTGGCTCGGCTCTGGACATTGACGAGCA 2625
QY 329 ----- 329
Db 2626 GAGTACATCAGATGGGGGCAAGGTGCCCATCAAGTGGAGTGGCTGGAGTCATTCTC 2685
QY 329 ----- 329
Db 2686 CGCGGGGTTCCACCACAGAGTATGTGAGTATGTGTACTGTGTGGAGCTG 2745
QY 329 ----- 329
Db 2746 ATGACTTTTGGGGCCAAACCTTACGATGGGATCCAGCCCGGGAATCCCTGACCTGCTG 2805
QY 329 ----- 329
Db 2806 GAAAAGGGGAGCGGCTGCCAGCCCATCTGACCATTTGATGTACATGATCATG 2865
QY 329 ----- 329
Db 2866 GTCAAAATTTGATGATTTGACTGTGAATGTGGCAAGATTCGGGAGTTGGTGTGAA 2925
QY 329 ----- 329
Db 2926 TTCTCCCGATGGCCAGGAGCCCGCTTTGTGTCATCCAGATGAGACTTGGGC 2985
QY 329 ----- 329
Db 2986 CCAGCAGTCCCTGGAGACACCTTTCACGCTCACTGTGAGAGAGATGACATGGG 3045
QY 329 ----- 329
Db 3046 GACCTGTGTGATGCTGAGAGATGTGTGATCCAGCAGGGCTTCTGTGTCAGACCT 3105
QY 330 ----- 347
Db 3106 GCTGGGAGCTCACAATAAGGCTGGAGCCCTTGAAGAGGAGGCCCGAGTCTCCAGT 3165
QY 348 GtlylAspLeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeu 367
Db 3166 GGTGGGAGCTCACAATAAGGCTGGAGCCCTTGAAGAGGAGGCCCGAGTCTCCAGT 3225
QY 368 AlaprosGluGlyAlaGlySerAspAlaPheaspGlyAspLeuGlyMetGlyAla 387
Db 3226 GCACCTCCGAAGGGGCTGGCTCCGATGTATTTGATGTGAGACTGTGGAATGGGGGCGCC 3285

Accession	Gene	Protein	Length (aa)	Score	Percent Similarity	Best Local Similarity	Query Match	DB
QY388	Lysylglycylserine	Prothrin	407	1437	64.93%	61.74%	58.53%	10
Db3286	MAAGGGCTGC	AAAGGCTCTCCACACATGACCCGACCCCTTACACCGGTACGTAGAGAC	334	426	64.93%	61.74%	58.53%	10
QY408	Prothrin	Protein Ser-Glu	427	1437	64.93%	61.74%	58.53%	10
Db3346	CCACAGTACCCCTGCTCCCTGACACATGATGGTACGTACGTGCCCCCTGACCTGCAGCC	340	426	64.93%	61.74%	58.53%	58.53%	10
QY428	Gln-Pro-Glu	Prothrin	447	1437	64.93%	61.74%	58.53%	10
Db3406	CAGCTTGAATATGTAACACACACATGTTGGCCCCAGCCCCCTTGCCCCCGAGAGGGC	346	426	64.93%	61.74%	58.53%	58.53%	10
QY448	Protein	Prothrin	467	1437	64.93%	61.74%	58.53%	10
Db3466	CCTCTGCTGCTGCCCCGACCTGTGTGTCACCTCTGGAAGAGCCCAAGACCTCTCCCA	352	426	64.93%	61.74%	58.53%	58.53%	10
QY468	Glycylserine	Prothrin	487	1437	64.93%	61.74%	58.53%	10
Db3526	GGGAAGATGGGGTGTCTAAAGAGCTTTTCCCTTTGGGGGTGCGCGAGAAACCCGAC	358	426	64.93%	61.74%	58.53%	58.53%	10
QY488	Tyrosine	Prothrin	507	1437	64.93%	61.74%	58.53%	10
Db3586	TACTTGACACCCCAAGGAGAGCTGCCCTGAGCCCAACCTCTCTGCTTCAGCCCA	364	426	64.93%	61.74%	58.53%	58.53%	10
QY508	Ala	Prothrin	527	1437	64.93%	61.74%	58.53%	10
Db3646	GCTTCGCAACCTTATCTGTGGACACGACACACAGAGCGGGGGCTCCACCCAGC	370	426	64.93%	61.74%	58.53%	58.53%	10
QY528	Thiophene	Prothrin	546	1437	64.93%	61.74%	58.53%	10
Db3706	ACCTTCAAGGAGACACCTACGCGACAGACACCAAGTACCTGGGTCTGGACGTGCCA	3762	426	64.93%	61.74%	58.53%	58.53%	10
RESULT 14								
US-09-821-883-10								
Sequence 10, Application US/09821883								
Patent No. US20020061310A1								
GENERAL INFORMATION:								
APPLICANT: Laus, Reiner								
APPLICANT: Vidovic, Damlir								
APPLICANT: Graddis, Thomas								
TITLE OF INVENTION: Compositions and Methods for Dendritic								
TITLE OF INVENTION: Cell-Based Immunotherapy								
FILE REFERENCE: 7636-0022.30								
CURRENT APPLICATION NUMBER: US/09/821,883								
CURRENT FILING DATE: 2001-03-30								
PRIOR APPLICATION NUMBER: US 60/193,504								
PRIOR FILING DATE: 2000-03-30								
NUMBER OF SEQ ID NOS: 30								
SOFTWARE: FastSeq for Windows Version 4.0								
SEQ ID NO 10								
LENGTH: 1437								
TYPE: DNA								
ORGANISM: Artificial Sequence								
FEATURE:								
OTHER INFORMATION: HER300*RG-M-CSF construct								
US-09-821-883-10								
Alignment Scores:								
Pred. No.:	6,896-181	Length:	1437					
Score:	2188.50	Matches:	426					
Percent Similarity:	64.93%	Conservative:	22					
Best Local Similarity:	61.74%	Mismatches:	31					
Query Match:	58.53%	Indels:	211					
DB:	10	Gaps:	4					
US-09-821-883-2 (1-690) x US-09-821-883-10 (1-1437)								
QY1	Met	Prothrin	20	1437	64.93%	61.74%	58.53%	10
Db1	ATGAGAGCTGACCCCTCTCTGCGCAGGCGACCAAGCCCTTACGCTTCTTCTTT	60						

Dd	61	CTGGCTTTTCTTGCGCTAGACGCGAAGTGTACTAGCCAAAGAGATTGGCGCGCGGGCGCGG	120
Oy	41	SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr	60
Dd	121	TGCGACCAAGTGTGCACCGCGACAGCATGAAGCTGGCGGCTCCCTGGCACTCCCGAACC	180
Oy	61	HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu	80
Dd	181	CACCTGGACATGCTCCGCGCACCTTACCAGGGCTGCCAGGTGGTGGCAAGGAAACTGGAA	240
Oy	81	LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspLleGlnGluValGln	100
Dd	241	CTCACTTACCTGGCCACCACCAAGCCAGGCTGCTCTCTCCAGCATATCCAGGAGGTGCAG	300
Oy	101	GlyTyrValLeuLleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAspGlyLe	120
Dd	301	GGCTACGGCTCATGTGCTCAACACCAAGTAGAGAGGTCCCACTGGCAGAGGCTTCGGATT	360
Oy	121	ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp	140
Dd	361	GTGGCAGGCACCCAGCTCTTTGAGAGCAACTATGCTCCGTGGCGCGGTGTACCAATTGGAGAC	420
Oy	141	ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln	160
Dd	421	CCGGCAACCAATTACCACCCCTGTACAGGGGCTGCCACAGAGGCTTCGCGGAGCTGCAG	480
Oy	161	LeuArgSerLeuThrGlnLleLeuLysGlyValLeuLleGlnArgAsnProGlnLeu	180
Dd	481	CTTGCAACCTCACAGAGATCTTGAAAGAGGGGTCTTGATCCAGCGGAACCCCAAGCTC	540
Oy	181	CysTyrGlnAspThrLleLeuTrpLysAspLlePheHisLysAsnAsnGlnLeuAlaLeu	200
Dd	541	TGCTAACCGAGCAGATTCTTGAGAGACATCTTCCACAAGAACCAACAGCTGGCTCTC	600
Oy	201	ThrLeuLleAspThrAsnArgSerArgAlaCysHisProCysSerPromeCysLysGly	220
Dd	601	ACACTGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCGGATGTGTAAAGGC	660
Oy	221	SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla	240
Dd	661	TCCCGCTCTCGGGAGAGAGATTCTGAGAGATTGTCAAGGCTGACCGCCACCTGCTGTGGC	720
Oy	241	GlyGlyCysAlaIleArgCysLysGlyProLeuProThrAspCysCysHisGlnCysAla	260
Dd	721	GGTGGCTGTGGCCCTGCAAGGGGCCACTGCCCACTAGCTGTGCCATGGACAGTGGCT	780
Oy	261	AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer	280
Dd	781	GCCGGCTGCACGGGCCCCAACCACTGTGAGCTGCTGGCTGCCCTCCACTCAACCAAGT	840
Oy	281	GlyLleCysGluLeuHisCysProAlaLeuValThrTyrTrsAsnThrAspThrPheGluSer	300
Dd	841	GGCACTGTGTGGCTGCACTGCCACGCCCTGGTCACTCAACAACAGACAGTTTGATGCC	900
Oy	301	MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr	320
Dd	901	ATGGCCATCCCGAGGGCGGCTATACATCTGGGGCCAGAGTGTGTGATGCTGCTGCCCTAC	960
Oy	321	AsnTyrLeuSerThrAspValGlySerGlyAlaGlyLysIleValHisHisArgHisAlaArg	340
Dd	961	AACATCCCTTTTACGGACAGTGGGATCCGCTAGC-----	993
Oy	341	SerSerSerThrArgSerGlyGlyLysAspLeuThrLeuGlnLeuGlnProSerGluGlu	360
Dd	993	-----	993
Oy	361	GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly	380
Dd	993	-----	993
Oy	381	AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro	400
Dd	993	-----	993


```
Db 395 GACAAATTCGCCCGCTCCACCCAGGAGAACCCAGAGGGGGCGGAGCTGCACACTT 454
OY 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAspProGlnLeuCys 181
Db 455 CGAAGTCTCAGAGATCCCTGAAAGGAGAGATTGTATCCGTGGGAACCTCAGCTTGC 514
OY 182 TyGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
Db 515 TACAGAGACATAGTTGTGTGAAGAGACGCTCTCCGCAAGATATACCAACGCGCTCGTC 574
OY 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysGlySer 221
Db 575 GATATAGACACCAATCGTTCCCGGGCTGTCACACTTGTGCCCCCGCTCAAGACAAAT 634
OY 222 ArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValLysAlaGly 241
Db 635 CACGTGTTGGGTGAGATCCGGAAAGACTGTCAGATCTTGACTGGCACCACTGTACAGT 694
OY 242 GlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAlaAla 261
Db 695 GGTGTGTCGGGTCAAGAGCGCGCTGCCACAGTGAAGTGAAGAGAGTGGCGCA 754
OY 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281
Db 755 GGGTGCACGGGGCCCAAGACATCTGACTGCGCTGCTCCCACTTCAATCATAGTGT 814
OY 282 IleCysGlnLeuHisCysProAlaLeuValThrTyraAsnThrAspThrPheGlnSerMet 301
Db 815 ATCTGTAGCTGCACCTCCAGCCCTGTCACCTACAAACAGACACCTTTGATGCTCATG 874
OY 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321
Db 875 CACAACTCAGAGGTCCCTACACTTGTGTGCCAGCTGCGTACACCTGCCCTTACAAAC 934
OY 322 TyrlleuSerThrAspValGlySer----- 329
Db 935 TACCTGTCTACGGAGAGTGGATCTGCACATCTGTGTGTCGCCGAATTAACCAAGAGGTC 994
OY 329 ----- 329
Db 995 ACAAGTGAAGACGGAACACAGCGTTTGAGAAATGCAGCAAGCCCTGTGCTCGAGTGC 1054
OY 329 ----- 329
Db 1055 TATGCTCTGGGCATGAGACACTTCGAGGGCGAGGCCATCACCACTGACAAATGTCAG 1114
OY 329 ----- 329
Db 1115 GAGTTTATGCTGCAGAAAGATCTTTGGAGCCCTGGCATTTTGGCCGAGAGCTTTGAT 1174
OY 329 ----- 329
Db 1175 GGGAGACCCCTCTCCGGCATTTGCTCCGCTGAGGCGTGAAGCTCCAAGTTCGAAACC 1234
OY 329 ----- 329
Db 1235 CTGAGAGATACAGGTACTGTACATCTCAGCATGCGCAGACAGTCTCCGAGACTTC 1294
OY 329 ----- 329
Db 1295 AGTGTCTTCAGAAACCTTCGAATCATTCGGGGAGCGATTCACGATGGCGGCTACTCA 1354
OY 329 ----- 329
Db 1355 TTGACACTGCAAGGCCCTGGGGATCCACTCGCTGGGGCTCGCTCTACTTCGGGAGCTGG 1414
OY 329 ----- 329
Db 1415 AGTGAATTGGCTCTGATTCACCGCAAGCCCATCTGTTGTGACACTGATACCTGG 1474
OY 329 ----- 329
Db 1475 GACCAGCTCTTCGGGAACCAATCAGAGCCCTGCTCCACAGTGGGAACCGCGCGGAAG 1534
OY 329 ----- 329
Db 1535 GACTTGTGCGTTCGAGAGCGCTTGTGTATCTACTGATGTGTGCCAGGGCTACTGCTGG 1594
OY 329 ----- 329
Db 1595 GGGCCAGGGCCCAACCACTGTGTCMACTGCAGTCATTTCTTGGGGCCACAGAGTGTGTG 1654
OY 329 ----- 329
Db 1655 GAGAGTGCAGAGTATGGAAGGGGCTCCCGGAGTATGTGAGTACAAAGCCGTGTG 1714
OY 329 ----- 329
Db 1715 CCGTGTACCCCGAGTGTCAAGCTCAAAACAGCTCAGAGACGTGCTTTGATGGAGGCT 1774
OY 329 ----- 329
Db 1775 GATCAGTGTGACAGCTGCGGCCACTACAGAGACTCGTCTGTGTGCTCGCTGCCCC 1834
OY 329 ----- 329
Db 1835 AGTGTGTGAACCGGACCTCTCTCTACATGCCATCTGGAAGTACCGGATGAGAGGGC 1894
OY 329 ----- 329
Db 1895 ATATGCCAGCGGTGCCCCCATCACTGACACCACTGCTGTGTGATGTGATGAAGAGGC 1954
OY 329 ----- 329
Db 1955 TGCCCAAGAGAGAGAGAGAGAGCCGCGTGATTCATTCATTCGCACTGAGAGGGCTC 2014
OY 329 ----- 329
Db 2015 CTGCTGTCTGATCTTATGTGTGTGTGTGTAATCTTAATCAACGAGAGACAGAGAG 2074
OY 329 ----- 329
Db 2075 ATCCGAAATATAGATGCGTAGGCTGCTCGAGAAACTGAGTTAGAGCCGCTGAGC 2134
OY 329 ----- 329
Db 2135 CCCAGCGAGCAATGCCCAACAGGCTCAGATCGGATCTTAAAGAGAGAGCTTAAGG 2194
OY 329 ----- 329
Db 2195 AAGGTAAAGTGTGATCAGAGACTTTTGGCACTGTCTACAAAGGCAATCTGGATCCCA 2254
OY 329 ----- 329
Db 2255 GATGGGAGAAATGGAATAATCCCGGTGCTATCAAGGTGTTGAGAGAAACACATCTCCT 2314
OY 329 ----- 329
Db 2315 AAAGCCAAAGAAATTTAGATGAAGCCTATGTGCTGTGCTGTCTCCGTAT 2374
OY 329 ----- 329
Db 2375 GTGTCCCGCTCCTGGGCATCTGCTGACATCCACAGTACAGCTGGTGGACAGCTTATG 2434
OY 329 ----- 329
Db 2435 CCCTAGCGCTGCTCTGTGACATGTCCGAAACACCGAGGTGGCTAGGCTCCAGAGAC 2494
OY 329 ----- 329
Db 2495 CTGCTCAACTGTGTGTTCAGATTGCCAAGGGATGAGTACCTGGAGAGCGTGGGCTT 2554
OY 329 ----- 329
Db 2555 GTACACAGGACCTGTGCTGCCGGAATGTGTAGTCAAGAGTCCCAACACGTCAGAGATT 2614
```

OY	329	-----	329	-----	329
Db	2615	ACAGATTTGGGGCTGGCTCGGCTGCGTGGACATTGATGACAGACAGTACCATGCAATGACG	2674		
OY	329	-----	329	-----	329
Db	2675	GGCAAGGTGGCCATCAAAATGGATGGCATTTGGAATCATTTCTGAGACGGCGGTCAACCAT	2734		
OY	329	-----	329	-----	329
Db	2735	CAGAGTATGTGTGGACCTATGGAGTGAAGTGTGTGGAGCTGATACATTTTGGGCCAAA	2794		
OY	329	-----	329	-----	329
Db	2795	CCTTACGATGGAAATCCAGCCCGGAGAGATCCCTGATTTCTGAGAGAGGAGACGCTTA	2854		
OY	329	-----	329	-----	329
Db	2855	CCTGAGCCTCCAAATCTGCACCAATTGATGTCATCATGATTAATGTCAAATGTTGATGATT	2914		
OY	329	-----	329	-----	329
Db	2915	GACTCTGAATGTGCCCCGAGATTCCGGAGATTGGTGTCAGAATTTTTCACGATTTGGCGAGG	2974		
OY	329	-----	329	-----	329
Db	2975	GACCCCGAGCGTTTGTGTGATTCACGAACGAGAGACTTGGGCCCATCCAGCCCATGGAC	3034		
OY	329	-----	329	-----	329
Db	3035	AGTACCTTACCGCTCAGTCGTGGAAGATGATGACATGGTGGACCTGGTACGCGTGA	3094		
OY	330	-----	330	-----	330
Db	3095	GAGTATCTGGTGGCCCGACGAGGATTTCTTCCCGGACCTTACCCAGCAGCACTGGAGC	3154		
OY	334	MetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyLysPheLeuThrLeu	353		
Db	3155	ACAGCCCATAGAAAGGCAACCCACACTGCTCCACCGAGAGTGGAGTGTAGCTGACACTG	3214		
OY	334	GlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGluAla	373		
Db	3215	GGCTGGAGGCGCTTGGAGAAAGGGCCCCCAGATTCCTCCACTGGCTTCCCTGGAGAGGGCT	3274		
OY	374	GlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeu	393		
Db	3275	GGCTTCGATGTGTGATGTGGAGCTGGCAATGGGGGTAAACAAAGGGCTGCAGAGCGTC	3334		
OY	334	ProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeuPro	413		
Db	3335	TCTCCACATGACCTCACGCCCTCTACACGCGGTACAGGAGAGACCCACATTAACCTCGACC	3394		
OY	414	SerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGluTyrValAsn	433		
Db	3395	CCCGAGACTGATGGCTATGTTGGCTCCCTGGCTGAGGCCCGACCGCAGATATGTGAAC	3454		
OY	434	GlnProAspValArgProGluProProSerProArgGluGlyProLeuProAlaAlaArg	453		
Db	3455	CAATCAGAGGTTCCAGCCTCAGCCCTTTAAACCCAGAGGGTCTCTGCTCTCTCCGG	3514		
OY	454	ProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGlyAlaVal	473		
Db	3515	CCTGCTGGTGGTACTCTAGAAAGACCCAGACCTCTCTCTCTGGAGAAATGGGGTTGTC	3574		
OY	474	LysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrProGlnGly	493		
Db	3575	AAAGACGTTTTTGGCTTCGGGGGGGTGTGGAGAACCCGTAATTAATTAACCGAGAGA	3634		
OY	494	GlyAlaAlaProGluProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr	513		
Db	3635	GGCACTGCTCTCCGCCCAACCCCTCTCCGCTTACAGCCCAAGCCCTTGTACAAACCTCAT	3694		

[illegible]

Search completed: May 8, 2003, 14:57:36
Job time : 285 secs

THIS PAGE BLANK (USPTO)

Fri May 9 14:21:07 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 8, 2003, 12:34:18 ; Search time 79 Seconds
(without alignments)
2678.569 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARASISLGLF.....EPVDEGAPPPAAHHHHH 690

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPRO_pool/US09821883/runat_28042003.104750.14590/app.query.fasta.1.839
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rnt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 ecgn.1.1.61.0runat.28042003.104750.14590 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEDQUERY -NEG_SCORES=0 -MATT=LONGLOG -DEV=TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCITUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2423	64.8	4530	1	US-08-229-515A-9
2	2423	64.8	4530	1	US-08-645-865-9
3	2423	64.8	4530	1	US-09-167-322-4
4	2418	64.7	4473	2	US-09-048-804-1
5	2418	64.7	4473	4	US-09-056-105-26
6	2405.5	64.3	3768	2	US-08-625-101-1
7	2405.5	64.3	3768	2	US-08-356-786-1
8	2258.5	60.4	2385	2	US-09-146-283-3
9	2258.5	60.4	2385	3	US-08-579-823A-3
10	2258.5	60.4	2385	4	US-09-344-195-3
11	1968	52.6	3955	1	US-08-229-515A-14
12	1968	52.6	3955	1	US-08-645-865-14

13	1647	44.0	1872	3	US-08-422-108-2
14	1647	44.0	1872	4	US-08-422-734-2
15	770	20.6	1588	2	US-09-146-283-1
16	770	20.6	1588	3	US-08-579-823A-1
17	770	20.6	1588	4	US-09-344-195-1
18	725.5	19.4	1868	4	US-08-658-863B-1
19	725.5	19.4	1868	4	US-09-676-610B-26
20	724.5	19.4	1593	4	US-09-676-610B-25
21	724.5	19.4	5532	2	US-08-475-035-3
22	724.5	19.4	5532	4	US-09-676-610B-17
23	722.5	19.3	896	6	5200327-2
24	722.5	19.3	900	1	US-08-318-193-7
25	721	19.3	5484	4	US-09-633-580A-3
26	721	19.3	5501	1	US-08-484-438-1
27	721	19.3	5551	1	US-08-484-438-3
28	719	19.2	905	6	5200327-3
29	719	19.2	909	1	US-08-318-193-9
30	714.5	19.1	4905	1	US-07-978-895-3
31	714.5	19.1	4905	1	US-08-473-119-3
32	714.5	19.1	4905	2	US-08-475-352-3
33	714.5	19.1	4975	4	US-09-630-706-3
34	710	19.0	1958	4	US-09-570-454-1
35	693	18.5	822	3	US-08-463-318-69
36	693	18.5	822	3	US-08-468-609A-69
37	693	18.5	822	4	US-08-446-872A-69
38	693	18.5	822	4	US-08-762-227A-69
39	693	18.5	822	5	PCT-US95-01185-69
40	685.5	18.3	903	3	US-08-468-318-66
41	685.5	18.3	903	3	US-08-468-609A-66
42	685.5	18.3	903	4	US-08-446-872A-66
43	685.5	18.3	903	4	US-08-762-227A-66
44	685.5	18.3	903	5	PCT-US95-01185-66
45	677	18.1	392	1	US-08-318-193-1

ALIGNMENTS

RESULT 1
US-08-229-515A-9
Sequence 9, Application US/08229515A
Patent No. 5518865
GENERAL INFORMATION:
APPLICANT: RAZIIDDIN
TITLE OF INVENTION: EPRB2 PROMOTER BINDING PROTEIN IN
TITLE OF INVENTION: NEOPLASTIC DISEASE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESS: NEEDLE & ROSENBERG PC
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/229,515A
FILING DATE: 19 APR 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PERRMAN, DAVID G
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414, 608
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-688-0770
TELEFAX: 404-688-9880
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 4530 base pairs

Sequence 2, Appl1
Sequence 2, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Sequence 25, Appl1
Sequence 25, Appl1
Sequence 17, Appl1
Patent No. 5200327
Sequence 7, Appl1
Sequence 3, Appl1
Sequence 1, Appl1
Sequence 1, Appl1
Patent No. 5200327
Sequence 9, Appl1
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 3, Appl1
Sequence 1, Appl1
Sequence 69, Appl1
Sequence 69, Appl1
Sequence 69, Appl1
Sequence 69, Appl1
Sequence 66, Appl1
Sequence 66, Appl1
Sequence 66, Appl1
Sequence 66, Appl1
Sequence 1, Appl1

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
us-08-229-515A-9

Alignment Scores:

Pred. NO.:	6.58e-161	Length:	4530
Score:	2423.00	Matches:	524
Percent Similarity:	41.14%	Conservative:	5
Best Local Similarity:	40.75%	Mismatches:	21
Query Match:	64.80%	Indels:	736
DB:	1	Gaps:	4

US-09-821-883-2 (1-690) x US-08-229-515A-9 (1-4530)

```

OY 2 Arg1Aa1aProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
DB 94 CGCGCCCTCCAGCGGGGTCCAGCGCGGATGGCGCGGAGCCGCGAGCCGAGCCATG 153
OY 18 PheLeuPheLeuPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 37
DB 154 GAGCTGGCGGCTTGTGCGCTGG-----GGGCTCTCTCTGCGCCCTTGGCCCCC 204
OY 38 GlyAla1a1aSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
DB 205 GGAGCCCGGACACCCAGATGTGCACCGGACAGACATGAAAGCTGGGCTCCCTGCGCAGT 264
OY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrgInglyCysGlnValValGlnGly 77
DB 265 CCCGAGACCCACCTGAGACATGCTCGCCACCTTACCCAGGGCTGCGAGTGTGTCAGGGA 324
OY 78 AsnLeuGluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
DB 325 AACCTGGAACTCACCTTACCTGCCACCAATGCCAGCTGTCTTCTTCCAGATATCCAG 384
OY 98 GluValGlnGlyTyrrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArg 117
DB 385 GAGGTGCAGGGCTACGAGCTATGCTATGCTACAAACAGAGGAGGAGTCCACTGCGAGG 444
OY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAsp 137
DB 445 CTGCGGATGTGCGAGGACCCAGCTCTTGTAGGACAACTTGGCCCTGGCTGCTGAC 504
OY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
DB 505 AATGAGAGCCCGTGAAATATCACCCCTGTACAGGGGCTCCCGCGAGGCTGGG 564
OY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValIleGlnArgAsn 177
DB 565 GAGCTGCAGCTTCAAGGCTTCAAGATCTTGAAGAAGGGGCTTGTATCCAGGGGAC 624
OY 178 ProGlnLeuCysTyrgInAspThrIleLeuThrLysAspIlePheHisLysAsnGln 197
DB 625 CCCAGGCTCTGCTACAGGACAGATTTTGTGAAGGACATCTTCCAAACAAACACAG 684
OY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
DB 685 CTGCTCTCACTGATGATAGACACCAACCGCTCTGGGGCTCCACCCCTTCTCCGATG 744
OY 218 CysLysGlySerArgCysTyrrGlyGlnSerSerGlnuAspCysGlnSerLeuThrArgThr 237
DB 745 TGTAAAGGCTCCCGCTGTGGGAGAGATTTCTAGGATTTCTCAAGCTGACCGGCT 804
OY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGln 257
DB 805 GTCTGTGCGGCTGTGCTGTGCGGCTGCAAGGGGACATCGCCCATGTGCTGCAATAG 864
OY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
DB 865 CAGTGTCTGTGCGGCTGACAGGCGCCCAAGACATCTGACTCTGGGCTGCTCCACTTC 924
OY 278 AsnHisSerGlyIleCysGlnLeuHisCysArgProAlaLeuValThrTyrrAsnThrAspThr 297

```

```

DB 925 AACCAAGTGCATCTGTGAGCTGCACCTGCCAGCCCTGGTCCACTTACACAGACAGACG 984
OY 298 PheGluSerMetProAsnProGluGlyArgTyrrPheGlyAlaSerCysValThrAla 317
DB 985 TTGAGTCCATGCCCAATCCGAGGGCGGTATACATTCGGCGGCGAGCTGTGTACTGTCC 1044
OY 318 CysProTyrrAsnTyrrLeuSerThrAspValIlySer----- 329
DB 1045 TGTCCCTCAACTACCTTCTTACGAGAGCTGGATCTGCACCTCTGTGCCCCCTGCAC 1104
OY 329 ----- 329
DB 1105 AACCAAGGTGACAGACAGAGATGAAACACAGCGGTGTGAGAAAGTCAGCAGACGCCCTGT 1164
OY 329 ----- 329
DB 1165 GCCGAGTGTGCTATGTGCTGGGATGAGGACATTGCGAGAGGTGAGGGCAGTTACAGCT 1224
OY 329 ----- 329
DB 1225 GCCAATATCCAGAGATTGTGCTGTGCTGCAAGAAGATCTTTGGAGGCTGCAATTTCTGCCG 1284
OY 329 ----- 329
DB 1285 GAGAGCTTTGATGGGGACCCAGCCCTCCAAACTGCCCCGCTTCAGGCCAGACAGCTCCAA 1344
OY 329 ----- 329
DB 1345 GTGTTGAGACTGTGAAGAGATCAGAGTTACGTATACATCTGAGATGSCCGGACAGC 1404
OY 329 ----- 329
DB 1405 CTGCTGACCTACAGCGTCTTCCAGAACCTGCAGATATCCGGGAGCAATTTCTGCACAAAT 1464
OY 329 ----- 329
DB 1465 GCGGCTTACTGCTGACCTTCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCTCACTG 1524
OY 329 ----- 329
DB 1525 AGGGAATGGGAGTGGAGTGGCCCTCATCCACATTAACACCACTTGTGCTGTGCAC 1584
OY 329 ----- 329
DB 1585 ACGGTGCCCTGGAGCAGCTCTTTGGAACCGGACCAAGCTGTCTCCACACTGCCAAC 1644
OY 329 ----- 329
DB 1645 CGGCCAGAGACGAGTGTGTGGCGAGGGCTTGGCTTGCACAGCTGTGCGCCGAGGG 1704
OY 329 ----- 329
DB 1705 CACTGTGGGGTCCAGGGGCCACCCAGTGTGTCAAGTGCACAGATTCCTTCGGGGCCAG 1764
OY 329 ----- 329
DB 1765 GAGTGCGTGAGGAATGCCAGTACTGCAAGGGGCTCCCGAGGAGTATGTGAAGCCAGG 1824
OY 329 ----- 329
DB 1825 CACTGTGTGCCGTGCCACCTTGAAGTGTACGCCCAAGATGGCTAGTACCTGTATTGGA 1884
OY 329 ----- 329
DB 1885 CCGGAGGCTGACAGTGTGTGCTGTGCGCCACTATTAAGACCTTCCCTTGTGCGTGGCC 1944
OY 329 ----- 329
DB 1945 CGTGTCCCAAGCGGTGTGAACCTGACCTTCTTACATGCCCCATGTGAAGTTTCCAGAT 2004
OY 329 ----- 329
DB 2005 GAGAGGGGCGCATGCCAGCTTGGCCCATGCACTGACACCCCACTCTGTGTGGAACCTGGAT 2064

```


329 ----- 329
Db 1765 GAGTCGTGAGAGATGCCGAGTACTGACAGGGGCTCCCGAGGAGTATGTAATCAG 1824
329 ----- 329
Db 1825 CACTGTTTGGCGTGCCACCCTGAGTGTACGCCAGAAATGGCTCAGTACCTGTTTGA 1884
329 ----- 329
Db 1885 CCGAGGCTGACCACTGTGTGCTGTGCCACTATTAAGAACCTTCCCTTGTGCTGGCC 1944
329 ----- 329
Db 1945 CGCTGCCCGAGGCGTGTGAACCTGACCTTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2004
329 ----- 329
Db 2005 GAGGAGGGCGCATGCCAGCCTTGCCCATCAACTGCACCCTCCTGTGTGAGACTGGAT 2064
329 ----- 329
Db 2065 GACAAGGGCTGCCCGCGGAGAGAGACCACTCTGAGCTCCATGCTCTGCGGTG 2124
329 ----- 329
Db 2125 GTTGCATTTCTGCTGTGCTGTGGGGGTGTCTTTGGGATCCTCATCAAGCGAGCG 2184
329 ----- 329
Db 2185 CAGCAGAAAGATCCGGAGTACACGATGCGGAGACTGTGCAGAAACGAGCTGGTGAG 2244
329 ----- 329
Db 2245 CCGCTGACACCTAGCGGAGCGGATGCCCAACGAGCGCAGATCCCTGGAAGAGAG 2304
329 ----- 329
Db 2305 GAGCTGAGGAAGTGAAGTGTCTTGATCTGCGCTTTTGACACAGTCTACAAGGCAATC 2364
329 ----- 329
Db 2365 TGGATCCCTGATGGGAGAAATGTAATAATCCAGTGGCCATCAAAAGTGTGAGGAAAC 2424
329 ----- 329
Db 2425 ACATCCCCCAAGCCAAACAATAATCTTAGACGAAGCATACGTATGCTGTGGGC 2484
329 ----- 329
Db 2485 TCCCCATATGTCTCCGCTTCTGGGCAATCTGACATTCACGAGTGACAGCTGTGACA 2544
329 ----- 329
Db 2545 CAGCTTATGCTTATGGCTGCTCTTAGACCATGTCCGGGAAACCGCGAGCGCTGGGC 2604
329 ----- 329
Db 2605 TCCCAGACCTGCTGACACTGTGTATGCAGATTGCCAAGGGGATGAGTACCTGGAGAT 2664
329 ----- 329
Db 2665 GTGCGGCTGTACACAGGAGATTGGCGGCTCGGAAGCTGTGTAAGATCCCAACAT 2724
329 ----- 329
Db 2725 GTCAAAATTACAGACTTGGGCTGCGCTGCTGTGACATTGACAGAGACAGATACAT 2784
329 ----- 329
Db 2785 GCAGATGGGGGCAAGGTGCCCATCAAGTGAATGGCGCTGGAGTCCATTCTCCGCGGCGG 2844
329 ----- 329

2845 TTCAACCACAGAGTATGTGTGAGTATGTTGACTGTGTGGAGCTGATGACTTTT 2904
329 ----- 329
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCTGACTGCTGAAAAAGGGG 2964
329 ----- 329
Db 2965 GAGCGGCTGCCCGACCCCGCATCTGCACCATATGATGTCTACATGATGATGTCATAATGT 3024
329 ----- 329
Db 3025 TGGATGATTGACTGTGAATGTGCGCAAGATTCCGGAGTGTGTGTGAATTCTCCGCG 3084
329 ----- 329
Db 3085 ATGCCAGGAGACCCCGAGCGCTTTGTGTATCCAGAAATGAGACTGGGCCACCGCAGT 3144
329 ----- 329
Db 3145 CCCTTGACAGACCTTCTACCCGCTGCTGAGAGACATGACATGGGGAGCTGGTG 3204
330 ----- gly 330
Db 3205 GATGCTGAGAGTATCTGTATCCCGAGCGGCTTCTCTGTCCAGACCCTGCCCCGGCG 3264
331 AlaGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyLys 350
325 GCTGGGCATGTGTCACACAGGACCGACACTCATCTACACAGGAGTGGCGGGAGC 3324
351 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 370
325 CTGACACTAGGGCTGAGCCCTCTGACAGAGAGGCCCCAGGTCTCCACTGGCACCTCC 3384
371 GluGlyAlaGlySerAspValPheAspLysPheGlyMetGlyAlaAlaLysGlyLeu 390
385 GAAGGGCTGGCTCCGATGATTTGATGTGCTGACCTGGGAATGGGGGAGCCAAAGGGCTG 3444
391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410
3445 CAAAGCTTCCCGACACATGATCCCGAGCTCTACAGCGGTACAGTGAAGAGACCCACAGTA 3504
411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
3505 CCCCCTGCTGACATGATGCTACGTGGCCCCCTGACCTGCAGCCCCCAGCTCTGACT 3564
431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyProLeuPro 450
3565 TATGTGAACACAGCATGCTTGGCCCCAGCCCCCTTGGCCCCGAGAGGGCTCTGCT 3624
451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaGlyAlaLysThrLeuSerProGlyLysAsn 470
3625 GCTGCCGACCTGTGTGGCCACTGTGGAAGAGGCCAAGACTCTTCCCGAGGAGAAAT 3684
471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
3685 GGGGTGCTCAAGAGCTTTTCCCTTGGGGGTGCGCTGGAGAACCCGAGACTTGGACA 3744
491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
3745 CCCCAGGGAGAGCTGCCCTCAGCCCAACCTCTCTGCTGAGCCACACTTTCGAC 3804
511 AsnLeuTyrTyrTrpAspGlnAspProProGluAlaValAlaProProSerThrPheLys 530
3805 AACCTCTATTAATGGGACAGGACCCACAGCGGGGGGCTCCACCGACCTTCAAA 3864
531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaPro 550
3865 GGGACACTACGGCAGAGAACCAAGTACCTGGGTCTGGAGCTGCCA-----GTGGA 3918
551 AlaArgSerProSerPro 556
|||||

```

DB 3919 ACCAGAGGCCAAGTCCG 3936

RESULT 3
US-09-167-322-4
: Sequence 4, Application US/09167322
: Patent No. 6365151
GENERAL INFORMATION:
APPLICANT: Allegheny University of the Health
Sciences, Halpern, Michael S.
England, James M.
TITLE OF INVENTION: CANCER VACCINE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seidel, Gonde, Lavorana & Monaco, P.C.
STREET: Suite 1800, Two Penn Center Plaza
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/167,322
FILING DATE: 07-Oct-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/00582
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 7933-33 PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4530 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-167-322-4

Alignment Scores:
Pred. NO.: 6,58e-161 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64.80% Indels: 736
DB: 4 Gaps: 4

US-09-821-883-2 (1-630) x US-09-167-322-4 (1-4530)
OY 2 ArgAlaAlaPProLeuLeuAlaArgA-----AlaSerLeuSerLeuGly 17
||||| ||| ::::|::|::| |||::|::|::|
DB 94 GCGGCCCTCCCAAGCCGGGTCCAGCCGGAGCCATGCGGCCGGAGCCGACGTAGACCATG 155
||||| ||| ::::|::|::| |||::|::|::|
OY 18 PheLeuPheLeuLeuPhePheTrPLeuAspArgSerValLeuAlaIaLysGluLeuAlaArg 37
||| ||| ||| ::::|::|::| |||
DB 154 GAGCTGGGGCGGCTTGCCGCTGG-----GGGCTCCTCCTCGGCCCTTGCCGCC 204
||||| ||| ||| ::::|::|::| |||
OY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuAlaGluProAlaSer 57
||||| ||| ||| ::::|::|::| |||
DB 205 GAGAGCGGGACACCCCAAGTGTGACCGGCACACACATGAAGCTGCGGCTCCCGCCAGT 264
||||| ||| ||| ::::|::|::| |||
OY 58 ProGluTrpHisLeuAspMetLeuArgHisLeuTyrgInglyCysGlnValIaGIngly 77
||||| ||| ||| ::::|::|::| |||
DB 265 CCGAGAGACCACCTGTGACATGCTCCGCCACCTCTACCAAGGCTCCAGGTGTGTCCAGGGA 322
||||| ||| ||| ::::|::|::| |||

```

QY	78	AsnLeuGluLeuThrTrpTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	325	AAACCTGGAACTCCACCTACTCTCCCAACAAATGCCAGCCTGTCTTCTCTCCAGAGTATCCAG	384
QY	98	GluValGlnGlyTyrValLeuIleAlaIlnAsnGlnValArgGlnValProLeuGlnArg	117
Db	385	GAGGTCCAGGGCTCACTGCTCATGCTCCACAAACCAAGTGGAGGAGCTCCACATGCCAGAG	444
QY	118	LeuArgIleValArgGlyThrGlnLeuPheGlnAspAsnTrpAlaLeuValIleLeuAsp	137
Db	445	CTGGGGATTGGCCAGAGACCACCACTCTTTGAGACAACTATCCCTGGCCGGCTGAGAC	504
QY	138	AsnGlyAspProLeuAsnAsnTrpThrProValTrpGlyValaSerProGlyLeuAsnArg	157
Db	505	AATGGAGACCCGGTGAACAAATACCAACCCCTGTACAGGGGGCTCCCAAGAGGCTGGGG	564
QY	158	GluLeuGlnLeuArgSerLeuThrGlnIleLeuValGlyGlyValLeuIleGlnArgAsn	177
Db	565	GAGCTGCAGCTTCCAAAGCGCTCACAGAACTTTGAAAGAGGGGTCTTGATTCACAGCGAAC	624
QY	178	ProGlnLeuCysTrpGlnAspThrIleLeuTrpAspIlePheIlnIstAsnAsnGln	197
Db	625	CCSCAGCTCTGTCAACAGACACATTTTGTGAAAGACATCTTCCAAAGAAACAAACAG	684
QY	198	LeuAlaLeuThrLeuIleAspTrpIleAsnArgSerArgAlaCysIlnAspCysSerProMet	217
Db	685	CTGGCTCTCACACGTAAGAAAGAACCAACCGCTCTCGGGCTGCCAACCCCTGTTCTCCGAA	744
QY	218	CysIstGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgTrpThr	237
Db	745	TGTAAAGGCTCCCGCTGCTGGGGAGAAAGATTGTGAGATGTGTACAGACCTGACGCCACT	804
QY	238	ValCysAlaGlyGlyCysValaArgCysValGlyProLeuProThrAspCysCysIstGlu	257
Db	805	GCTCTTGCCGGGTGGCTGGCCCGCTCGAAAGGGGCCACTGCCCTCACTGCTCTCCACTGAG	864
QY	258	GlnCysAlaAlaGluCysTrpGlnProIstGlnIleSerAspCysLeuAlaCysLeuIlnAspHe	277
Db	865	CAGTGTGCTGGCGGCTGGACGGGGCCCAAGCACTGTACCTGGCTGGCTGCTCCACTTTC	924
QY	278	AsnIlnSerGlyIleCysGlnLeuIlnIstCysProAlaLeuValIlnTrpTrpIlnThrAspThr	297
Db	925	AAACCACTGGCCATCTGAGACTSCACTGCTCCCAAGCCTGTGTCACTAACAACAGACAGC	984
QY	298	PheGluSerMetProAsnProGlnGluArgTrpThrPheGlyAlaSerCysValaIlnTrpAla	317
Db	985	TTTGATGTCACAGCCCAATCCCGAAGGGCGGATACATTCGGCCCAAGCTGTGTGACTGTCC	1044
QY	318	CysProTrpAsnTrpLeuSerThrAspValAlaIleSer	329
Db	1045	TGTCCCTCAACTACTCTTCTTACGGAGTGGGATCTCGAACCTCGTCTGGCCCCCGTGCAC	1109
QY	329	-----	329
Db	1105	AAACAAGAGTGCACAGAGAGATGGAACAACAGCGGTGTGAGAACTGCACAAGCCCTGT	1166
QY	329	-----	329
Db	1165	GCCCAAGTGTCTATGTGTCTGGCATGTAGCACTTGCAGAGGTGAGGGCAGTATCAAGT	1222
QY	329	-----	329
Db	1225	GCCAAATATCCAGAGATTGTGTGGCTGCAMAAGATCTTTTGGAGACCTGCATTTTGGCG	1284
QY	329	-----	329
Db	1285	GAGAGCTTTGATGGGAGACCCAGCCTCCAAACACTGCCCCCTCCAGCCAGAGCACTCCAA	1344
QY	329	-----	329
Db	1345	GTGTTTGAAGACTCTGGAAAGAGATCACAGTTTAACTATACATCTCACACATGGGGCGGACAC	1400
QY	329	-----	329

Db	1405	CTGGCTGACCTCACACGCTCTTCCAGAAACCTGCAAGTAACTCCGGGGAGCAAAATCTGCACAA	1464
0Y	329	-----	329
Db	1465	GGCCGCTACTGCTGACACCTCGCAAGGGCTTGGGCATCAGCTGGCTGGGCTGGCTCACTG	1524
0Y	329	-----	329
Db	1555	AGGGAACCTGGGCACTGGACTGGCCCTCAATCCACCATAAACCCACCTCTGCTTGGTGAC	1584
0Y	329	-----	329
Db	1585	ACGGTGGCCTTGGGACACGCTCTTTGGGAACCCGCAACAAGCTCTGCTCCACACTGCCAAC	1644
0Y	329	-----	329
Db	1645	CGGCCAGAGACAGAGTGTGTGGCGAGGGCTGGCTGGCACAGCTGTGCGCCGAGGG	1704
0Y	329	-----	329
Db	1705	CACGTCTGGGGTCCAGGGCCCAACCAAGTGTCACTGCAAGCAAGTCTTCGGGGCCAG	1764
0Y	329	-----	329
Db	1765	GAGTCCGTGGAGGAATCCCGAGTACTGTCAGAGGGCTCCCGAGGAGTATGTGAATGCCAAG	1824
0Y	329	-----	329
Db	1825	CACGTGTTGCCGTGCCACACCTGAGTGTCAAGCCCCAAGATGGCTCAAGTACGCTTTTGA	1884
0Y	329	-----	329
Db	1885	CCGGAGGCTGACAGTGTGTGGCTGTGCCACTATAAGAGACCCTCCCTTCTGCTGGCC	1944
0Y	329	-----	329
Db	1945	CGCTGCCACAGCGGTGTGAACCTGACCTCTCTACATGCCATCTGGAATTTCCAGAT	2004
0Y	329	-----	329
Db	2005	GAGGAGGGCGCATCCAGACCTTGGCCCATCACTGACACCACCTCCTGTGTGACCTGGAT	2064
0Y	329	-----	329
Db	2065	GACAAAGGCTGCCCCGCCGAGACAGAGACGAGCCTCTGACGTCCATGCTCTCGGGTG	2124
0Y	329	-----	329
Db	2125	GTTGGCATTTCTGTGTCGTGTGTCTTGGGGGTGTCCTTTGGGATCCTCATCAAGGACGG	2184
0Y	329	-----	329
Db	2185	CAGCAGAAGATCCGGAAGTACAGATGCGGAGACTCTCGAGGAAAGGAGCTGGGAG	2244
0Y	329	-----	329
Db	2245	CCGCTGACACTAGCGGAGCGATGCCAACAGAGCGCATGCGGATCCTGAAGAGAGC	2304
0Y	329	-----	329
Db	2305	GAGCTAGAGAGGTGAAGGTGCTTGGATCTGGCGCTTTTGGCAGACTCTAAGGGGATC	2364
0Y	329	-----	329
Db	2365	TGGATCCTGATGGGGGGAATGTGAATAATTCAGTGGCCATCAAAAGTGTGAAGGAAAC	2424
0Y	329	-----	329
Db	2425	ACATCCCCAAAGCCACAAGAAATCTTAGACGAAGATACGTGATGCTGTGTGGC	2484
0Y	329	-----	329

Db	2485	TTCCCATATGCTCTCCGCGCTTCTGGGCACTGTGCCTGACATCCAGAGTGCACCTGGTGCA	2544
QY	329	-----	329
Db	2545	CAGCTTATGCCCTATGGCTGCCTTTAGACATGTCCGGAAACC CGGACGCTGGGC	2604
QY	329	-----	329
Db	2605	TTCCAGGACCTGCTGAACCTGCTGTATGCAAGATTGCCAAGGGAGTACCTACCTGGAGAT	2664
QY	329	-----	329
Db	2665	GTGGGGCTCGTACACAGGAGCTTGGCGCGCTCGGAACGTGCTGTCAAGATCCCAACAT	2724
QY	329	-----	329
Db	2725	GTCAAATATACAGACTTCGGGCTGCCTCGGCTGTGACATTTGACGACAGATACCAT	2784
QY	329	-----	329
Db	2785	GCAAGTGGGGGCAAGGTGCCCATCAAGTGAGTGGCGCTGGAGTCCATTCTCCCGGCGG	2844
QY	329	-----	329
Db	2845	TTCAACCAACCAAGATGATGTGTGAGTTTATGATGTACCTGTGGGACTATGACTTTT	2904
QY	329	-----	329
Db	2905	GGGGCAACACTTACGATGGGATCCAGCCCGGAGATCCTGACCTGCTGAAAAAGG	2964
QY	329	-----	329
Db	2965	GAGCGGCTGCCAGAGCCGCCCATCTGACCAATTGATGTCACATGATCATGGTCAATGT	3024
QY	329	-----	329
Db	3025	TGGATGATTGACTCTGAATGTGGGCCAAGATTCCGGGAGTTGGTGTCTGAATTCTCCGC	3084
QY	329	-----	329
Db	3085	ATGGCAGAGGACCCCAAGCGCTTTGTGTGCATCCAGAAATGAGGACTTGGGCCAGCAGT	3144
QY	329	-----	329
Db	3145	CCCTTGGACAGACCTTCTACCGCTCACTGTGAGAGACGATGACATGGGGACCTGGTG	3204
QY	330	-----Gly 330	
Db	3205	GATGCTGAGAGATCTCTGTGACCCCAAGAGGCTTCTGTCCAGACCCTGGCCCCGGGC	3264
QY	331	AlaGlyIymeValIhSIhSIATgSIhSIATgSerSerSerThiArgSerGlyIyIyAsp	3350
Db	3265	GCTGGGGGCAATGGTCCACACAGGACCGGAGTCACTACAGGAGTGGGGTGGGAGC	3324
QY	331	LeuThiLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer	370
Db	3325	CTGACACTAGAGGCTGGAGCCCTCTGAAGAGAGGCCCCAGGCTCTCCACTGGCACCTCC	3384
QY	371	GluGlyIhIaGlySerAspValIhIhAspGlyIyAspLeuGlyIymeGlyAlaIhIyGlyLeu	390
Db	3385	GAAAGGGCTGGCTCCGATGTATTGTAGTGGACTGGGAATGGGGGACCCAAAGGGGCTG	3444
QY	391	GlnSerLeuProThiIhSIAspProSerProLeuGlnArgTyrSerGluAspProThiVal	410
Db	3445	CAAAAGCTCCCAACACTGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCCAACATA	3504
QY	411	ProLeuProSerGluThiAspGlyTyrValAlaProLeuThiCysSerProGlnProGlu	430
Db	3505	CCCCGCTCTGTGAGATGTATGGCTACAGTTGGCCCCCTGACCTGCAGCCCCAGCCTGAA	3564
QY	431	TyrValaIhGlnProAspValIhIhArgProGlnProProSerProArgGluGlyProLeuPro	450
Db	3565	TATGTGAACCAAGCAGATGTGTGGCCCCAGCCCCCTTGCCCCCGAGAGGGCCCTCTGCT	3624

QY 451 AlAlaArgrProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn 470
 Db 3625 GCGGCCGACCTGCTGTGGCAGCTCTGAAAGGCCAAGACTCTCTCCAGGGAAGAAAT 3664
 QY 471 GlyAlaValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGlyTyrLeuThr 490
 Db 3685 GGGGTCTCTCAAGACGCTTTTGGCTTGGGGGTCCCTGAGAACCCCGAGTACTTACA 3744
 QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
 Db 3745 CCCAGGAGAGAGCTGCCCTCAGCCCCCAGCCCTCTCTCTCAGCCCAAGCCTTGAC 3804
 QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
 Db 3805 AACCTCTATTACTGGACACGAGACCCACAGAGCGGGGCTCCACCCAGCAGCTTCAAA 3864
 QY 531 GlyThrProThrAlaGluAsnProGlyTyrLeuGlyLeuAspValProAlaAlaAlaPro 550
 Db 3865 GGGACACTTACGGCAGAGAACCCAGAGTACCTGGTCTGGACGTGCA-----GTGTGA 3918
 QY 551 AlaArgSerProSerPro 556
 Db 3919 ACCAGAGGCCAGAGTCCG 3936
 RESULT 4
 US-09-048-804-1
 ; Sequence 1, Application US/09048804
 ; Patent No. 5968748
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett, Allan Lipton, Lois M. Witters
 ; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE MODULATION OF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5968748rls LLP
 ; STREET: One Liberty Place - 46th Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: U.S.A.
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 1.44 Mb diskette
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/048,804
 ; FILING DATE: Herewith
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Paul K. Legard
 ; REGISTRATION NUMBER: 38,534
 ; REFERENCE/DOCKET NUMBER: ISIS-2913
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4473 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Unknown
 ; ANTI-SENSE: NO
 ; US-09-048-804-1
 Alignment Scores:
 Pred. No.: 1.45e-160 Length: 4473
 Score: 2418.00 Matches: 523
 Percent Similarity: 41.06% Conservative: 5
 Best Local Similarity: 40.67% Mismatches: 22
 Query Match: 64.67% Gaps: 736
 DB: 2

US-09-821-883-2 (1-690) x US-09-048-804-1 (1-4473)
 QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
 Db 118 CGGCCCTCCACCGCGGGTCCAGCCGAGCATATGGGCGCGGAGCGGAGAGACCATTG 177
 QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGlyLeuAlaArg 37
 Db 178 GAGTGGCGGGCTTGTGCGCGTGG-----GGGTCTCTCTCGCCCTTGGCCCCCCC 228
 QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
 Db 229 GGAGCCCGGAGACCCCAAGTGTGCACCGGACAGACACTGAACTGCGCGCTCCCTCCAGT 288
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
 Db 289 CCCGAGACCCACCTGGACATATGCTCCGACCTCTACACAGGGCTGCGAGTGTCCAGGGA 348
 QY 78 AsnLeuGluLeuThrTyrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
 Db 349 AACCTGGAACCTACACCTACCTGCGCCACCAATGCCAGCTGTCTCTCTGAGAGATATCCAG 408
 QY 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
 Db 409 GAGGTGAGGGGCTACGCTCATCTGCTCACACCAAGTGAAGGAGGTCCGACCTGAGAGG 468
 QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
 Db 469 CTGGGATTTGTGCGAGGACCCACCTCTTTGAGGACACATATGCTCCGCGCGCTGTAGAC 528
 QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
 Db 529 AATGAGACACCCGCTGAAACATATACACCCCTGTACAGAGGCGCTCCACGAGAGGCTCGG 588
 QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsn 177
 Db 589 GAGTGCAGCTTGAGAACCTTCACAGAGATCTTGAAAGAGGGGCTTGATCCAGCGGAC 648
 QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGln 197
 Db 649 CCCAGCTCTGTACACAGAGACAGATTTTGTGAAGAGACATCTTCCACAGAACACACAG 708
 QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
 Db 709 CTGGCTCTCACACTGATAGACACCAACCGCTCTGCGGCTGCGCACCCCTGTCTCCGATG 768
 QY 218 CysLysGlySerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThr 237
 Db 769 TGTAAAGGCTCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAAGCCTGACGCGCAGCT 828
 QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGln 257
 Db 829 GTCTGTCGCGGTGGCTGCGCGCTGCAAGGGGCACTGGCCACTGACTGCTGCGCAATGAG 888
 QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
 Db 889 CAGTGTCTCCGCGCTGCAAGGCGGCCCAACACCTGACTGCTGCGGCTGCTGCTGCTGCTG 948
 QY 278 AsnHisSerGlyLysGlyGlyLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
 Db 949 AACCAACAGTGCATCTGTGAGCTGCACTGCAAGCCTGTGTCACTCAACACAGACAGCAG 1008
 QY 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
 Db 1009 TTGTAGTCCATGCCCAATCCCGAGGGCGGATATATCATTCGGCGGCACTGTGTGACTGCC 1068
 QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
 Db 1069 TGTCCCTACAAATACCTTTCTACGAGAGTGGGATTCCTGCACACCTCGTCTGCCCTGAC 1128
 QY 329 ----- 329

Db 1129 AACCAAGAGTACAGACAGAGGATGAAACACAGCGGTGTGAGAAATGCAGCAAGCCCTGT 1188
OY 329 ----- 329
Db 1189 GCCCGAGTGTCTATGTGTGGGCGATGAGACACTTTCGAGAGGTGAGGGCAGTTACCAGT 1248
OY 329 ----- 329
Db 1249 GCCAATATCCAGAGATTGTGCTGCGTCAAGAAGATCTTTGGAGCGCTGGCATTTCTGCGC 1308
OY 329 ----- 329
Db 1309 GAGAGCTTTGATGGGAGCCAGACCTCCAACACTGCCCGCTCCAGCAGAGAGCTCCAA 1368
OY 329 ----- 329
Db 1369 GTGTTGAGACTCTGGAAAGATCAGAGTTACCTATATCTCAGCATGCGCCGACAGC 1428
OY 329 ----- 329
Db 1429 CTGCTGACCTCAGCGTCTTCAGAACTGCAGATTAATCCGGGGAGCAATTTCTGCACAAT 1488
OY 329 ----- 329
Db 1489 GGGCCCTACTGCTGACCTCGCAAGGGCTGGGCGATCAGCTGGGCTGGGCTGAGCTG 1548
OY 329 ----- 329
Db 1549 AGGGAAGTGGGAGTGGAGCTGCGCTCATCCACATTAACACCACCTCTGCTTCTGAC 1608
OY 329 ----- 329
Db 1609 ACGTGCCCTGGGAGCACACTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAAC 1668
OY 329 ----- 329
Db 1669 CGGCGAGAGAGAGTGTGTGGGCGAGGGCGTGGCTGCCACAGCTGTGCGCCGAGGG 1728
OY 329 ----- 329
Db 1729 CACTGCTGGGGTTCAGGGGCCACCCAGTGTGTCAACTGAGCGCACTTCTTGGGGCCAG 1788
OY 329 ----- 329
Db 1789 GAGTGGAGAGATGCCAGTACTGCAAGGGGCTCCGAGGAGTATGTGAATGCCAGG 1848
OY 329 ----- 329
Db 1849 CACTGTTGCCCTGCCACCTGAGTGTAGCCCCAGAAATGGCTCAGTACCTGTTTGGGA 1908
OY 329 ----- 329
Db 1909 CCGGAGGCTGACCAAGTGTGTGCGCTGTGCCACTATTAAGAGACCTCCCTTCTGCGTGGCC 1968
OY 329 ----- 329
Db 1969 CGCTCCCCAGGCGTGTGAACCTGACCTCTCTACATGCCCCATGTGAAGTTTCCAGAT 2028
OY 329 ----- 329
Db 2029 GAGGAGGCGCATGCGAGCTTGCCCATCAACTGACCACTCCTGTGTGAGACTGGAT 2088
OY 329 ----- 329
Db 2089 GACAAGGCTGCCCGCGAGCAGAGAGCCAGCCCTGTGAGCTCATCATCTCTGCGGTG 2148
OY 329 ----- 329
Db 2149 GTTGCAATTTCTGCTGTGTGTCTTTGGGGGTGTGCTTTGGGATCCTCATCAAGGAGCG 2208
OY 329 ----- 329
Db 2209 CAGCAGAAGATCCGGAAGTACAGATGCGAGACTGTGCAAGAAACGAGCTGGTGGAG 2268

OY 329 ----- 329
Db 2269 CCGCTGACACCTAGCGAGGCGATGCCAACAGGCGGAGATGCGGATCTGAAAGAGACG 2328
OY 329 ----- 329
Db 2329 GAGCTGAGGAAGTGAAGTCTTGATCTGGCGCTTTGGCACAGTCTACAGGCGATC 2388
OY 329 ----- 329
Db 2389 TGGATCCCTGATGGGAGGAATGTGAAAAATTCAGTGGCCATCAAGTGTGAGGAAAAAC 2448
OY 329 ----- 329
Db 2449 ACATCCCCCAAGCCACAAAGAAATCTTAGACGAGACATACGTATGCTGTGTGGGC 2508
OY 329 ----- 329
Db 2509 TCCCATATGTCTCCGCGCTTCTGGGCACTGCTGACATCCAGGTCACACTGTTGACA 2568
OY 329 ----- 329
Db 2569 CAGCTTATGCCCTATAGCTGCTCTCTAGACCATGTCCGGAAAAACCGGAGCGCTGGGC 2628
OY 329 ----- 329
Db 2629 TCCAGACCTGTGACTGTGTATGCAATTTGCCAAGGGGATGAGCTTACTGAGAGAT 2688
OY 329 ----- 329
Db 2689 GTGGGCTGTACACAGGAGTGTGGCGCTCGAAGCGTGTGTCAGAGTCCCAACCAT 2748
OY 329 ----- 329
Db 2749 GTCAAAATTACAGACTTGGGCTGGCTGCTGTGACATTGACGAGACAGATACCAT 2808
OY 329 ----- 329
Db 2809 GCAGATGGGGGCAAGTCCCATCAAGTGAATGGGCGTGGAGTCCATTCTCGCGCGCGG 2868
OY 329 ----- 329
Db 2869 TTCAACCCAGAGTGAATGTGAGATTATGTGACTGTGTGGAGCTGATGACTTTT 2928
OY 329 ----- 329
Db 2929 GGGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCTGACCTGCTGAAAAGGGG 2988
OY 329 ----- 329
Db 2989 GAGCGGCTGCCCCAGCCCCCATCTGCACATTTGATGTCTACATGATCATGGTCAATGT 3048
OY 329 ----- 329
Db 3049 TGGATGATTGACTGTGAATGTGCGCCAGAAATTCGCGAGTTGTGTAATTTCCCGC 3108
OY 329 ----- 329
Db 3109 ATGGCAGGAGACCCCAAGCGCTTGTGTGTCATCCAGAAATGAGGACTTGGGCCAGCCAGT 3168
OY 329 ----- 329
Db 3169 CCCTTGGACAGACCTTCTACCGCTCACTGCTGGAGAGCATGACATGGGAGACTGTG 3228
OY 329 ----- 329
Db 3229 GATGCTGAGAGTATCTGTACCCAGAGGCGCTTCTGTCCAGACCTTGGCCCCGGGGC 3288
OY 329 ----- 329
Db 3289 GCTGGGCGATGTCTCACACAGAGCCGACGAGCTCATCTTACCAAGGAGTGGCGGTGGGAC 3348
OY 329 ----- 329

Db	1129	AACCAAGAGGTGACAGCAGAGATGGAAACACAGCGGTGTGAAATCTGCACGACAGCCCTGT	1188
0y	329	-----	329
Db	1189	GCCCGAGTGTCTATGGTCTGGGCATGAGACACTTGCAGAGGTGAGGGCATTACAGT	1248
0y	329	-----	329
Db	1249	GCCAAATTCACAGACGTTTGCTGCTGCACAGAAAGATCTTTGGAGCCTTGACATTTCTGCCG	1308
0y	329	-----	329
Db	1309	GAGAGCTTTGATGGGAGCCACAGCCTCCAAACACTGGCCCCGCTCCAGCAGAGAGCTCCAA	1368
0y	329	-----	329
Db	1369	GTTGTTGAGACTGTGAGAGATCACAGGTTACTATACATCTCAGCATGGCCGACAGC	1428
0y	329	-----	329
Db	1489	GGCGCTACTCGCTGACCCCTGCAGAGGGCTGGGCATCAGCTGGCGGGCTCGCTCACTG	1548
0y	329	-----	329
Db	1549	AGGAACTGGGCACTGGACATGGACCTGCGCCCTCATCCACCATAACCCACCTCTGCTTGTCAC	1608
0y	329	-----	329
Db	1609	ACGGTGCCCTGGGACACAGCTCTTTCGGAACCGCACCAAGTCTGTCTCCACACTGCGCAAC	1668
0y	329	-----	329
Db	1669	CGGCCAGAGCAGAGTGTGTGGCGAGAGGCTGGCCTCCACACAGTGTGGCCCGAGGG	1728
0y	329	-----	329
Db	1729	CACGTGTGGGATCCAGGGCCCAACCCAGTGTGTCAACTGCAGCCAGCACTTCCTTCGGGGCAG	1788
0y	329	-----	329
Db	1789	GAGTGCTGGAGAGATCCGAGTACTGACAGGGGCTCCAGAGGATATGTAAATCCAGG	1848
0y	329	-----	329
Db	1849	CACGTTTGGCGGTCCACCTGTGAGTGTCAAGCCCAAGATGGCTCAGTACCTGTTTGGAA	1908
0y	329	-----	329
Db	1909	CCGAGAGCTGACAGTGTGTGGCCTGTGCCACTATAAGACCCTCCCTTCGCGTGCC	1968
0y	329	-----	329
Db	1969	CGCTGCCCAAGCGGTGTGAAACTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT	2028
0y	329	-----	329
Db	2029	GAGGAGGCGCATCCAGCCTTGCCCATCACTGCAACCCACTCTGTTGTGAGACTGGAT	2088
0y	329	-----	329
Db	2089	GACAAAGGCTGCCCGCCGAGCAGAGAGCCAGCCTTGACGTCCATCATCTTCCGGGTG	2148
0y	329	-----	329
Db	2149	GTTGGCAATTCTGCTGTGCTGTGCTTGGGGGTGTCTTTGGGATCCTCATCAAGCAGCG	2208
0y	329	-----	329
Db	2209	CAGCAGAAGATCCGAGATACAGATCGGAGACTGCTCAGAAACCGAGCTGGTAG	2268

[illegible]

|||||
Db 466 CACGGGAACCCAGCTCTCTACAGAGACACGATTTTGGAGAGACATCTCCACAG 525
QY 195 AAsAsnGlnLeuAlaLeuThrLeuLeuIleAspThrAsnArgSerArgAlaCysHisProCys 214
Db 526 AACACCAAGCTGGCTCTCAACACTGATAGACACCAACCGCTCTGGGGCTCCACCCCTGT 585
QY 215 SerProMetCysLysGlySerArgCysTTPGlyGlySerSerGlnuAspCysGlnSerLeu 234
Db 586 TCTCCGATGTGTAAAGGCTCCCGCTGCTGGGGAGAGAGTTCTAGAGATTCTCAGAGCTTG 645
QY 235 ThrArgThrValCysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCys 254
Db 646 ACCGCACTGTCTGTGGCTGGCTGTGCCCGCTGCAGAGGGCCACTGCCACTGACTGC 705
QY 255 CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys 274
Db 706 TGGCATGAGCAGTGTGTGGCGGCTGCACGGGCCCAAGCACTCTGACTCTGGGCTGC 755
QY 275 LeuHisPheAsnHisSerGlyIleCysGlyLeuHisCysProAlaLeuValThrTyrAsn 294
Db 766 CTCGACTTCAACCAAGAGTGGCATCTGTGAGCTGCACTGCCAGCCCTGTCTACCTACAC 825
QY 295 ThrAspThrPheGlnSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCys 314
Db 826 ACAGACACGTTGAGTGCATGCCCCAATCCGAGGGCCGATACATTCGGCGCAGCTGT 885
QY 315 ValThrAlaCysProTyrAsnTyrLeuSerThrAspAlaGlySer 329
Db 886 GTGACTGCTGTCCCTACAACTCTTCTACGAGCTGGGATCTGCACCTCTGCTGC 945
QY 329 329 329
Db 946 CCCCTGCACAAACAGAGGTGACAGACAGAGATGGAACACAGCGGTGTGAAAGTGCAGC 1005
QY 329 329 329
Db 1006 AAGCCCTGTGCCGAGTGTGCTATGCTGTGGCATGTGAGCATTGGCAGAGTGAAGGCA 1065
QY 329 329 329
Db 1066 GTTACCAAGTCCCAATATCAGAGAGTTTCTGTGCTGCAAGAAGATCTTTGGAGCTGCA 1125
QY 329 329 329
Db 1126 TTCTCTGCCGAGAGCTTGTGATGGGGACCAAGCTCCACACTGCCCGCTCCAGCCAGAG 1185
QY 329 329 329
Db 1186 CAGCTCCAAGTGTGTAGACTCTGGAGAGATACAGATTACCTATACATCTCAGCATGG 1245
QY 329 329 329
Db 1246 CCGAGACAGCTGCTGACTCAGAGCTTCCAGAACCTGCAGATTAATCCGGGAGCAATT 1305
QY 329 329 329
Db 1306 CTGACAAATGGCGCTACTGCTGACCTGCAGAGGCTGGGCATCAGCTGGCTGGGGCTG 1365
QY 329 329 329
Db 1366 CGCTCACTGAGAGAACTGGGCAGTGAGACTGGCCCTCATTCACACATTAACACCACTCTGC 1425
QY 329 329 329
Db 1426 TTGCTGACACAGGTGCTGGGACAGCTTTTCGGAACCGACCAAGCTGTGCTCCAC 1485
QY 329 329 329
Db 1486 ACTGCCAACGGGCCAGAGAGAGAGTGTGTGGGCGAGGGGCTTGGCTGCCACCAAGCTGTGC 1545
QY 329 329 329

Db 1546 GCCCGAGGCGACTGTGGGCTCCAGGGCCACCAGTGTGTCAACTGCAGCCAGTTCTT 1605
QY 329 329 329
Db 1606 CGGGGCCAGAGATGCGGTGGAGGAATGCCAGTACTGACAGGGGCTCCCGAGGAGTATGTG 1665
QY 329 329 329
Db 1666 AATGCCAGGACGTGTTCGGCTGCCACCCGATGAGTGTACGCCAGAAATGGCTCAGTGACC 1725
QY 329 329 329
Db 1726 TGTTTGGACCGGAGGCTGACAGATGTGTGCTGTGCCACTATAAGACCTCCCTTC 1785
QY 329 329 329
Db 1786 TGGGTGGCCCGCTGGCCCAAGCGTGTGAACCTGACCTCTCTACATGCCCATCTGGAG 1845
QY 329 329 329
Db 1846 TTTCCAGATGAGGAGGGCGCATGCCAGCTTGGCCCATCAACTGCACCCACTCTGTGTG 1905
QY 329 329 329
Db 1906 GACCTGGATGACAAAGGCTGCCCGCAGACAGAGACCAAGCCCTGACCTCATCATC 1965
QY 329 329 329
Db 1966 TCTGGGTGTGGCATTTGCTGTGCTGTGCTTGGGGTGTCTTTGGATCTTCATC 2025
QY 329 329 329
Db 2026 AAGCGACGCGACGAGAAAGATCCGGAAGTACACGATGGGAGACTGTGCAGAAACGGAG 2085
QY 329 329 329
Db 2086 CTGGTGAAGCCGCTGACACCTAGCGAGCGATGCCAACAGGCGCAGATCGGATCTGT 2145
QY 329 329 329
Db 2146 AAAGACAGGAGCTGAGAAAGTGAAGTCTTGATCTGGCGCTTTTGGACACATCTAC 2205
QY 329 329 329
Db 2206 AAGGCATCTGATCCCTGATGGGAGAAATGTGAAAATTCAGATGGCCATCAAAAGTGTG 2265
QY 329 329 329
Db 2266 AAGGAAACACATCCCCCAAGCCAAACAAAGAAATCTTAGACGAAAGCATACGTGATGGCT 2325
QY 329 329 329
Db 2326 GGTGTGGCTCCCAATATGTCTCCCGCTTCTGGGATCTGCTGACATCCAGCGTGAC 2385
QY 329 329 329
Db 2386 CTGGTGACACAGCTTATGCCCTATGSGCTGCTCTTAGACCATGTCCGGAAACCGGGGA 2445
QY 329 329 329
Db 2446 CGCCTGGGCTCCAGAGACTGTGAACTGTGTATGACAGATTCGCAAGGGGATGAGCTAC 2505
QY 329 329 329
Db 2506 CTGGAGATGTGGGCTGTCTACACAGGAGCTTGGCCGCTGGAAAGTGTCTGATACAGT 2565
QY 329 329 329
Db 2566 CCCAAGCATGTCAAAATTACAGACTTGGGGCTGGGCTGGGCTGTGACATTTGACGAGACA 2625
QY 329 329 329
Db 2626 GAGTACCATGCAGATGGGGCAAGGTGCCCATCAAGTGAATGAGCGCTGAGATCCATTCTC 2685

```

OY 329 ----- 329
Db 2686 CGCCGGGCTTACACCCAGAGTATGTGTGAGTATGTGTGACTGTGTGGAGCTG 2745
OY 329 ----- 329
Db 2746 ATGACTTTTGGGCAAACTTACAGATGGATCCAGCCGGAGATCCCTGACTGCTG 2805
OY 329 ----- 329
Db 2806 GAAAGGGGAGGGGCTCCCAAGCCCATCTGACCATTTGATGCTACATGATCATG 2865
OY 329 ----- 329
Db 2866 GTCAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2925
OY 329 ----- 329
Db 2926 TTTCTCCCGCATGCGCAGAGACCCCGCTTTGTGTGATCCAGATGAGACTGGGC 2985
OY 329 ----- 329
Db 2986 CAGCCAGTCCCTTGGACAGACCTTCTACCGCTGCTGAGAGACGATGACATGGG 3045
OY 329 ----- 329
Db 3046 GACCTGGTGGATCTGAGAGATCTGTGATCCAGAGGGCTTCTGTCCAGACCT 3105
OY 330 ----- 347
Db 3106 GCGCCGGGCTGAGGAGGCTGACACAGGACGCGAGCTCATCTACAGAGAGG 3165
OY 348 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3225
Db 3166 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3225
OY 368 AATPProSerLutLysLysLysLysLysLysLysLysLysLysLysLysLysLys 387
Db 3226 GACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3285
OY 388 LysGlyLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 407
Db 3286 AAGGGGCTGAAAGCCCTCCACACATGACCCACCTTACAGCGGTACAGAGAGC 3345
OY 408 ProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerPro 427
Db 3346 CCCACAGTACCCCTGCGCTGTAGAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAG 3405
OY 428 GlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGluGly 447
Db 3406 CAGCTGAAATGTGTGAACCGAGATGTGCGCCAGCCCTTCCGCGCCAGAGAGG 3465
OY 448 ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgLysThrLeuSerPro 467
Db 3466 CTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3525
OY 468 GlyLysAsnGlyValValLysAspValPheAlaPheGlyValAlaValGluAsnProGlu 487
Db 3526 GGGAGAGATGGGTGCGCAAGAGCTTTTGGCTTTGGGGGTGCGTGGAGAACCCGAG 3585
OY 488 TyrLeuThrProGlnGlyAlaAlaProGlnProHisProProAlaPheSerPro 507
Db 3586 TACTTACACCCAGAGAGAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3645
OY 508 AlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSer 527
Db 3646 GCGCTTGACAACTCTATTACTGTGGACAGAGCCACCAAGGGGGGCTCCACAGCAG 3705
OY 528 ThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValPro 546
Db 3706 ACCTTCAAGAGGAGACCTACAGGAGAGAACCCAGAGTACTGGGTCTGGAGCTGCCA 3762

```

```

RESULT 8
US-09-146-283-3
; Sequence 3, Application US/09146283
; Patent No. 5976546
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Compositions
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dellinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/146,283
; FILING DATE: 03-SEPT-1998
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; TELEPHONE/DOCKET INFORMATION: 7636-0010, 21
; TELEPHONE: 650-324-0860
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: homo sapiens
; INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
; US-09-146-283-3
Alignment Scores:
Pred. No.: 9,99e-150 Length: 2385
Score: 2258.50 Matches: 490
Percent Similarity: 64.598 Conservative: 19
Best Local Similarity: 62.188 Mismatches: 126
Query Match: 60.408 Indels: 154
DB: 2 Gaps: 19
US-09-821-883-2 (1-690) x US-09-146-283-3 (1-2385)
OY 35 LeuAlaArgGlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeu 54
Db 56 TTTCCCGCCCGGAGCCCGGAGACCCCAAGTGTGACCGGACAGACATGAAAGCTGGGCTC 115
OY 55 ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnVal 74
Db 116 CTTGACAGTCCGAGAGCCCACTGTGACATGCTCCGACACCTTACCAAGGGCTCCAGAGT 175
OY 75 ValGlnGlyAsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGln 94
Db 176 GTCCAGAGAAACCTTGAACTCACTACCTGCGCCACCAATGCCAGCCGTCTCTCCAGAG 235
OY 95 AspIleGlnGluValGlnGlyTyrValLeuIleuAlaHisAsnGlnValArgGlnValPro 114
Db 236 GATATCCAGAGAGGTCCAGGGCTACGTGCTCATGCTTCACAAACCAAGTGAAGAGGCTCCCA 295

```


Oy 115 LeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAla 134
 |||||
 Db 296 CTCAGAGAGGCTGGGATGTGGCGAGGACACCACTCTTTGAGAGCAACTATGCGCCGGCC 355
 Oy 135 ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGly 154
 |||||
 Db 356 GTGCTGAGCAATGTGAACCCGCGTGAACCAATACCACTCTGTCAAGGGGCTCCAGGA 415
 Oy 155 GlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIleuLeuArgGlyValLeuIle 174
 |||||
 Db 416 GGCCTCGGGAGCTGCAGCTTCGAGGCTCAGACAGATCTTGAAGAAGAGGGGCTTGATC 475
 Oy 175 GlnArgAspProGlnLeuGlyArgGlnAspThrIleLeuTrpLysAspIlePheHisLys 194
 |||||
 Db 476 CAGCGAAGCCCGAGCTGTGCTACAGAGACAGCATTTTGGAGAGGACATCTTCCACAG 535
 Oy 195 AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCys 214
 |||||
 Db 536 AACACACAGCTGGCTCTACACTGATAGACACCAACCGGCTCGGGCTGCCACCCCTGT 595
 Oy 215 SerProMetCysArgGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeu 234
 |||||
 Db 596 TCTCCATGTGTAAAGGCTCCGCTGCGGAGAGAGATTGTGAGATGTGTCAAGGCTG 655
 Oy 235 ThrArgThrValCysAlaGlyCysAlaArgCysLysGlyProLeuProThrAspCys 254
 |||||
 Db 656 AGCGGACGTGTGTGCTGCGGCTGTGCGCGCTGCAAGGGGCGCTGCCACATGACTGC 715
 Oy 255 CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys 274
 |||||
 Db 716 TGCACAGACACAGTGTCTGCGGCTGACAGGCGCCCAAGACACTGTACTGCGGCTGC 775
 Oy 275 LeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTrpAsn 294
 |||||
 Db 776 CTCACATTCACACACAGTGTGTGACTGTGACTGCTGACCCGCTGCTGCTGCTGCTG 835
 Oy 295 ThrAspThrPheGluSerMetProAsnProGluGlyArgTrpThrPheGlyAlaSerCys 314
 |||||
 Db 836 ACAGACACGTTGTGATGCCAATGCCCAATCCGAGGCGGATATCATTTGCGCGCACCTGT 895
 Oy 315 ValThrAlaCysProTrpAsnTrpLeuSerThrAspValGlySer----- 329
 |||||
 Db 896 GTGACGTGCTGTGCTCCACAACTATCTTCTACGAGCTGGAGTCTGCAACCCCTGCTGTGC 955
 Oy 329 ----- 329
 Db 956 CCCCTGCACACCAAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAACTGCAGC 1015
 Oy 330 -----GlyAlaGlyGlyMetValHisHisArgHisArgSer 341
 |||||
 Db 1016 AAGCCCTGTGCCGAGTGTGCTATGTGTG---GGCATGTGAGACACTTGGAGAGGTGAGG 1072
 Oy 342 SerSerThrArgSerGly-----GlyGlyAspLeu 351
 |||||
 Db 1073 GCAGTTACCACTGTCCAAATATCCAGAGATTGTGCTGCTGCAAGAAAGATCTTTGGGAGCTGT 1132
 Oy 352 ThrLeuGlyLeuGlnProSerGluGlnGluAlaProArgSerProLeuAlaProSerGlu 371
 |||||
 Db 1133 GCATTT---CTGCCGAGAGACTTGTGATGGGAGACCCCTCCAAACACTGCGCGCTCCAG 1189
 Oy 372 GlyAlaGlySerAspValPheAsp-----GlyAspLeuGlyMetGly 385
 |||||
 Db 1190 CCAGACACAGCTCCAAATGTTTGTAGACTGTGGAAGAGATCAGAGTTACCATATCATCTCA 1249
 Oy 386 AlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGln----- 402
 |||||
 Db 1250 GCATG---GCCGAGACAGCTCTGACTGACCTGAGCTCTTCCAGAACTGCAAGTATCCGGGG 1308
 Oy 403 -----ArgTrpSerGluAspPro----- 408
 |||||
 Db 1309 ACGAATTCCTGCACAAATGGCGCTACTGCTGACCCCTGCAAGGGCTGGCATCAGCTGCGCT 1368
 Oy 409 -----ThrValProLeuProSer----- 414

Db 1369 GGGCTGCGCTCAGTGAAGGAACTGGGCAAGTGAGTGGCCCTCATCATACCAATGAACCCA 1428
 Oy 415 -----GluThrAspGlyTrpValAlaProLeuThrCysSerProGlnProGlyTrp 431
 |||||
 Db 1429 CCTCTGCTCTGTCACACAGCTGGCTGGGAGCAGCTTTTGGAAACCCCGCAACAC--- 1485
 Oy 432 ValAsnGlnProAspValArgProGln-----ProProSer 443
 |||||
 Db 1486 ---TCTGCTCCACACTGCAACCGGCGCAGAGACAGACTGTGGGCGAGGCTGGCTGTG 1542
 Oy 444 ProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLys 463
 |||||
 Db 1543 CCA-----CGAGCTGTGCGCGCGAGGCGCACTGCTGGGTCAGAGGCGCCAC 1587
 Oy 464 ThrLeu-----SerProGlyLysAsnGlyVal----- 473
 |||||
 Db 1588 CCACTGTGTCAACTGCAGACAGTTCCTGCGGCGCGAGAGTGGCTGAGGAATGCGCAGT 1647
 Oy 474 -----LysAspValPheAlaPheGlyGlyAla 482
 |||||
 Db 1648 ACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGTTCCTGCCACCCCTGA 1707
 Oy 483 ValGluAsnProGlyTrpLeuThr-----ProGlnGlyGly 494
 |||||
 Db 1708 GTGTCAAGCCCAAAATAGGCTGACAGTGTTTTGGACCGAGGCTGACCAAGTGTGTGG 1767
 Oy 495 AlaAlaPro-GlnProHisPro-----ProProAlaPheSerPr 507
 |||||
 Db 1768 CTGTGCCCACTAATGAAGCACTCTTCTGCGGCGCGCTGCCCGCGGCTGTGAACCC 1827
 Oy 507 AlaPheAspAsnLeuTrpTrpAspGlnAspProProGluArgGlyAla---ProPr 526
 |||||
 Db 1828 TGACCTCTCTTACATGCTCCATCTG---AAGTTTCCAGATGAGAGGAGGCGCATGCCAGCC 1884
 Oy 526 oSerThrPhe-----LysGlyThrProThrAl 535
 |||||
 Db 1885 TTGCCCCATCAACTGCACCCACTCTGTGTGACACTGTGATGAACAAAGGGTGGCCCGCGA 1944
 Oy 535 AsnAsnProGlyTrpLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProse 555
 |||||
 Db 1945 GCAGAGAGCCAGCCCTGTG-----ACGTCCCTGAGAGCACCGCCGCTGCGCCAG 1995
 Oy 555 rProSerThrGlnProTrpGluHisValAsnAlaIleGlnGluAlaArgArgLeuAsn 575
 |||||
 Db 1996 CCCACACACAGCCCTGGAGAGCATGTGAATGCCATCCAGAGGCCCGGCTCTCTGAA 2055
 Oy 575 nLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIleSerGluMetPr 595
 |||||
 Db 2056 CTGAGTAGAGACACTGTGCTGAGATGAATGAACAGTAGAAGTCATCTCAGAAATGTT 2115
 Oy 595 AspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuTrpLysGlnGlyLeuAr 615
 |||||
 Db 2116 TGACCTCCAGAGAGCCAGCTGCTTACAGACCCCGCTGAGCTGTACAAACAGAGGCTGTGG 2175
 Oy 615 GGIysrLeuThrLysLeuLysGlyProLeuThrMetAlaSerHisTrpLysGlnHis 635
 |||||
 Db 2176 GGGCAGCTTCACCAAGCTCAAGGGCCCTTGACAGATGATGGCCAGGCACTACAAACAGCA 2235
 Oy 635 scysProProThrProGluThrSerCysAlaThrGlnIleIleThrPheGluSerPheLys 655
 |||||
 Db 2236 CTGCTCCACACCCCGGAACCTTCGTGCAACCCAGATTAATCACTTTGMAACTTTCAA 2295
 Oy 655 sGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGluProValGlnol 675
 |||||
 Db 2296 AGAGAACCTGAAGAGACTTCTGCTGTCAATCCCTTTGACTGTGTGGAGCCAGTCCAGGA 2355
 Oy 675 uGlyAlaProProProAla 682
 |||||
 Db 2356 GTGAGACCGGCCAGATGAGGCT 2377
 RESULT 9
 US-08-579-823A-3

```

: Sequence 3, Application US/08579823A
: Patent No. 6080409
: GENERAL INFORMATION:
: APPLICANT: Laus, Reiner
: APPLICANT: Ruegg, Curtis L.
: APPLICANT: Wu, Hongyu
: TITLE OF INVENTION: Immunostimulatory Composition and Method
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESS: Dehlinger & Associates
: STREET: 350 Cambridge Ave. Suite 250
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/579,823A
: FILING DATE: 03-DEC-1998
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Judge, Linda R.
: REGISTRATION NUMBER: 42,702
: REFERENCE/DOCKET NUMBER: 7636-0010
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-324-0980
: TELEFAX: 650-324-0960
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2385 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: homo sapiens
: INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
: US-08-579-823A-3

Alignment Scores:
Pred. No.: 9.99e-150 Length: 2385
Score: 2258.50 Matches: 490
Percent Similarity: 64.59% Conservative: 19
Best Local Similarity: 62.18% Mismatches: 126
Query Match: 60.40% Indels: 154
DB: 3 Gaps: 19

US-09-821-883-2 (1-690) x US-08-579-823A-3 (1-2385)
OY 35 LeuAlaArgIyAlaIaIaSerThrgInValIcysThrgIyThraSpMetLysLeuArgIeu 54
Db 56 TTGGCCCCCGGAGCCGCGAGCCAGTGTGCACCGCGCACAGACATGAACTCGCGCTC 115
OY 55 ProIaIaSerProgluThraIaIaSerMetLeuArgIhIaIaIaIaIaIaIaIaIaIa 74
Db 116 CTTGCGAGTCCCGAGACCCACCTGACATGCTCCGCGCACCTCTACCAAGGCTGCGAGGTG 175
OY 75 ValGlnGlnLeuGlnLeuThryrLeuProThraIaIaSerLeuSerPheLeuGln 94
Db 176 GTGAGGGAACCTGGAACCTACTACTGCTCCACACATGCGACACTCTCTCTCTGAG 235
OY 95 AspIleGlnGlnValGlnGlyTyrValLeuIleAlaIaIaIaIaIaIaIaIaIaIaIaIa 114
Db 236 GATATCCAGGAGGTGAGGCTGCTCATGCTCATGCTCAACCAAGTGAAGGAGTCCCA 295
OY 115 LeuGlnArgLeuArgIleValArgIyThrgInLeuPheGluAspAspTyrAlaLeuAla 134

```

```

Db 296 CTGCAGAGCGCTGCGAGATTGCGAGGACCCAGCTCTTTGAGCAACATATGCTGCGCC 355
OY 135 ValLeuAspAsnGlyAspProLeuAsnAsnThrProValThrgIyAlaIaSerProGly 154
Db 356 GTGCTAGACATGGAGACCCGCTGACAAATACACACCCCTGTACAGAGGCGCTCCCAAGA 415
OY 155 GlyLeuArgIleuGlnLeuArgSerLeuThrgIleuIleuLysGlyGlyValLeuIle 174
Db 416 GGCCTGCGGAGCTGCACCTTCGAAAGCTCAGACAGATCTTGAAGGAGGCGTCTTGATC 475
OY 175 GlnArgAsnProgluThryrGlnAspThrIleLeuThryrPheIleuAspIlePheHisLys 194
Db 476 CAGCGAAGCCCGAGCTCTGCTACAGACAGATTTTGTGAAGAGACATTTCCACAG 535
OY 195 AsnAsnGlnLeuAlaLeuThryrLeuIleAspThraIaIaSerArgAlaIaIaIaIaIaIa 214
Db 536 AACAAACAGCTGGCTGTACACTGTATGACACCAACCGCTCTCGGGCGTCCACCCCTGT 595
OY 215 SerProMetCysLysGlySerArgCysTrpGlyLeuSerSerGluAspCysGlnSerLeu 234
Db 596 TCTCCGATGTGTAAAGGCTCCCGCTGCGGAGAGAGACTTCTGAGGATGTCAAGAGCTG 655
OY 235 ThrArgThraValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThraIaIa 254
Db 656 AGCGGACAGTGTGTGTGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
OY 255 CysHsIaGlnGlnCysAlaIaIaGlyCysThrgIyProLysHisSerAspCysLeuAlaIa 274
Db 716 TGCCATGAGCAGTGTGTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCTG 775
OY 275 LeuHsIaPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaIleValThryrAsn 294
Db 776 CTCACCTTCAACACAGAGGATGTGTGAGCTGACCTCCAGCGCTGCTGCTGCTGCTGCTG 835
OY 295 ThrAspThrPheGluSerMetProAsnProgluIaArgIyThryrPheGlyAlaIaSerCys 314
Db 836 ACAGACAGCTTGTAGTCCATCCATCCCAATCCAGAGGCGGCTTATCATTTGCGCCAGCTGT 895
OY 315 ValThraIaCysProTyrAsnTyrLeuSerThraIaIaIaIaIaIaIaIaIaIaIaIaIa 329
Db 896 GTGACTGCTGCTGCTTACAACTTCTTACGAGCAGTGGATCTGTCACCTCGTCTGTC 955
OY 329 ----- 329
Db 956 CCCCTGCAACAAGAGTGTGACAGAGAGTGAACACAGCGGTGTGAGAACTGACAGC 1015
OY 330 -----GlyAlaGlyIleValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 341
Db 1016 AAGCCCTGTGCGCGAGCTGTCTATGTGCTG---GCCATGAGACACTTGGCGAGAGTGAAG 1072
OY 342 SerSerThraArgSerGly-----GlyAlaIaIaIaIaIaIaIaIaIaIaIaIaIa 351
Db 1073 GCAGTTACCACTGCCAATATTCACAGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1132
OY 352 ThrLeuGlyLeuGlnProSerGlnGlnGlnAlaProArgSerProLeuAlaProSerGln 371
Db 1133 GCATTT---CTGCGGAGAGAGCTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1189
OY 372 GlyAlaGlySerAspValPheAsp-----GlyAspLeuGlyIleCys 385
Db 1190 CCAAGAGAGCTCCAAAGTGTGTGAGACACTGTGAAGAGATCACAGGTTCATCATCTCA 1249
OY 386 AlaAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 402
Db 1250 GCATG-GCGGAGAGCGCTGACCTGACCTCAGCGTCTTCCAGAACTGCAAGTAATCCGGGG 1308
OY 403 -----ArgTyrSerGluAspPro----- 408
Db 1309 ACGAATTCGACAAATGCGCTTACTGCTGACCTTCACAAAGGCTGGGCAATCAGCTGCT 1368
OY 409 -----ThraValProLeuProSer----- 414
Db 1369 GGGGCTGCGCTCACTGAGGAGAACTGGGAGTGGAGTGGCTTCAATCCATCAACACCA 1428

```

```

Qy 415 -----glutrasp1tyr-ValAlaProLeuThrCysSerProGlnProGluTyr 431
Db 1429 CCTGCTGCTGTCACACAGGTCCTGGACACAGCTTTTGGAAACCCGACCAAGC--- 1485
Qy 432 ValAsnGlnProAspValArgProGln-----ProProSer 443
Db 1486 ---TCTGCTCCACACTGACCAACCGCCAGAGAGAGTGTGTGGCCAGGCGCTGGCCTG 1542
Qy 444 ProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluAlaAlaLys 463
Db 1543 CCA-----CCAGCTGTGGCCCGGAGGCGACTGCTGGGGTCCAGCGGCCAC 1587
Qy 464 ThrLeu-----SerProGlyLysAsnGlyValVal----- 473
Db 1588 CCACTGTGTCACTGACAGCCAGTCTCTGGGGCCAGAGAGTGGTGGAGAGATGCGAGT 1647
Qy 474 -----LysAspValPheAlaPheGlyGlyAla 482
Db 1648 ACTGACAGGCGCTCCCGAGGAGATGTGAATGCCAGAGCATGTTGGCCGACCCCTGA 1707
Qy 483 ValGluAsnProGluTyrLeuThr-----ProGlnGly 494
Db 1708 GTGTCAAGCCCGAGATGGCTGACCTGTTTGGACCGAGGCTGACAGTGTGTGGC 1767
Qy 495 AlaAlaPro-GlnProHisPro-----ProProAlaPheSerPr 507
Db 1768 CTGAGCCACATATAGAGCCCTCTCTGCTGGTGGCGCCGCTGCCAGCGGTGGAAC 1827
Qy 507 AlaPheAspAsnLeuTyrTyrPaspGlnAspProProGluArgGlyAla---ProPr 526
Db 1828 TGACCTCTCTACATGCTCCATCTG6---AGCTTCCAGATGAGGAGGCGCTCCACCC 1884
Qy 526 oSerThrPhe-----LysGlyThrProThrAl 535
Db 1885 TTGCCCATCACTGACCCACCTCTGTGTGACCTGGATGACAGGGCTGCCCGCGCA 1944
Qy 535 AsnAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSe 555
Db 1945 GCAGAGAGCCAGCCCTG-----ACGTCCCTGAGAGGACCCGCGCTGCCAG 1995
Qy 555 rProSerThrGlnProTrpGluHisValAsnAlaIleGlnGluAlaArgLeuLeuAs 575
Db 1996 CCCCAGGACACAGCCCTGGGAGCATGTGAATGCCATCCAGAGAGCGCCGCTCTCTGAA 2055
Qy 575 nLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIleSerGluMetPh 595
Db 2056 CCTGAGTAGACACACTGCTGTGAGATGAATGAACAGTAGACATCTCAGAAATGTT 2115
Qy 595 AspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuTyrLysGlnGlyLeuAr 615
Db 2116 TGACCTCCAGAGGCGGACCTCTCTACAGACCCGCTGAGGTGATCAAGAGCGGCTCG 2175
Qy 615 gGlySerLeuThrLysLeuLysGlyProLeuThrMetAlaSerHisTyrLysGlnH1 635
Db 2176 GGGGAGGCTTCAACCAAGGCGCCCTTGACCATGATGCGCACCACTCAAAACACCA 2235
Qy 635 sCyProProThrProGluTyrHisCysAlaThrGlnIleIleThrPheGluSerPheLy 655
Db 2236 CTGCCCTTCAACCCCGGAAACTTCTGTGCAACCCAGATATACACTTTGAAAGTTTAA 2295
Qy 655 sGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGluProValGlnG1 675
Db 2296 AGAGAACTGAAGAGCTTGTGCTGTGATCCCTTTGACGTGGAGGCCAGTCCAGGA 2355
Qy 675 uGlyAlaProProProAla 682
Db 2356 GTGAGAGCGGCGCATGAGGCT 2377

```

RESULT 10
 US-09-344-195-3
 ; Sequence 3, Application US/09344195
 ; Patent No. 6210662

```

GENERAL INFORMATION:
APPLICANT: Laus, Reiner
                Wu, Hongyu
                Ruegg, Curtis L.
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2385 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: GM-CSF-HER-2 fusion gene; Fig. 8
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-344-195-3
Alignment Scores:
Pred. No.: 9,99e-150 Length: 2385
Score: 2258.50 Matches: 490
Percent Similarity: 64.59% Conservative: 19
Best Local Similarity: 62.18% Mismatches: 126
Query Match: 60.40% Indels: 154
Gaps: 19
US-09-821-883-2 (1-690) x US-09-344-195-3 (1-2385)
Qy 35 LeuAlaArgGlyAlaAlaSerThrGlnValLcysThrGlyThrAspMetLysLeuArgLeu 54
Db 56 TTGCCCGGAGGAGCGGAGGAGCAACCAAGTGTGACCGGCGACAGACATGAACCTGGGCTC 115
Qy 55 ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnVal 74
Db 116 CCTGCAATCCCGAGAGACCACCTGACATGCTCGGCCACCTTACCAAGGCTGCCAGGTG 175
Qy 75 ValGlnGlyAsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGln 94
Db 176 GTGCAGGGAAACTGGAATCACTACCTGCCCAACATGCGACGCTGCTCTCTCGCAG 235
Qy 95 AspIleGlnGluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValPro 114
Db 236 GATATCCAGGAGGTGACAGGCTACGCTCATTCCTCAACCAAGATGAG-CAAGTCCCA 295

```

QY 115 LeuGlnArgLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAla 134
 DB 296 CTGAGAGGCTGGCGATGTTGGCGAGGCCAGCTCTTTGGAGCAACTATGCTGCTGGCC 355
 QY 135 ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGly 154
 DB 356 GTGCTTAACATATGAGAACCCGCTGACAAATATACACCCCTGTCACAGGGGCTCCCGAGA 415
 QY 155 GlyLeuArgGluLeuGlnLeuArgSerLeuThrGluIleLeuGlyGlyValLeuIle 174
 DB 416 GGCTGGGGAGCTGCAGCTTCGAAGCTTCAGAGATCTTGAAGAGAGGGGTCTTATC 475
 QY 175 GlnArgAsnProGlnLeuGlyTyrGlnAspThrIleLeuTryptylsAspIlePheHisLys 194
 DB 476 CAGGGAGACCCCGACGCTCTCTACAGACACGATTTTGGAGGACATCTTCCACAG 535
 QY 195 AsnAsnGlnLeuAlaLeuThrIleAspThrAsnArgSerArgAlaCysHisProCys 214
 DB 536 AACAAACAGCTGGCTCTCACACTGATGACACCAACCGCTCTGGGGCTGCCACCTGT 595
 QY 215 SerProMetCysLysGlySerArgCysTyrGlyLysLeuSerSerGluAspCysGlnSerLeu 234
 DB 596 TCTCCGATGTAGTAAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGATTTGTCAGAGCTG 655
 QY 235 ThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCys 254
 DB 656 ACGGCGACTCTCTGTGGGGTGGCTGTGCCCGCTGCAAGGGGCCACTGCCACTGACTGC 715
 QY 255 CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCys 274
 DB 716 TGCCTATAGCAAGTGTGCTGGCGGCTGCAGCGGCCCAAGCACTGTGACTGCTGGCTGC 775
 QY 275 LeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsn 294
 DB 776 CTCGCACTTCAACACACAGCTGATGCTGTGACTGCTGACCTGCCAGCCCTGGTCACTAC 835
 QY 295 ThrAspThrPheGluSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCys 314
 DB 836 ACGAGACAGTTTGTAGTCCATAGCCCAATCCGAGGGCCGGATATCATTCGGCGCCAGCTG 895
 QY 315 ValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySer 329
 DB 896 GTAGCTGCCGTCTCTCACTACACTTCTTACGAGACGTGGATCTCACCCTCTGCTGC 955
 QY 329 329
 DB 956 CCCCTGCACAACAGAGGTGACAGCAGAGATGGAACACAGCGGTGTGAAGATGCAGC 1015
 QY 330 330
 DB 1016 AAGCCCTGCCCCGAGTGTCTATGTCTG---GGCATGGAGCACTTGGCAGAGGTGAGG 1072
 QY 342 SerSerThrArgSerGly 1132
 DB 1073 GCAATTACAGTGCATATCCAGAGATTTGCTGGCTGCAGAAAGATCTTTGGAGGCTG 1132
 QY 352 ThrLeuGlyLeuGluProSerGluGlnAlaProArgSerProLeuAlaProSerGlu 371
 DB 1133 GCATTT---CTGCCGAGAGCTTTGATGGGAGCCCAACCTCCAAACACTGCCGCTCAG 1189
 QY 372 GlyAlaGlySerAspValPheAsp 385
 DB 1190 CCAAGACAGCTCCAGAGTGTGAACTGTGAAGAGATCAGACAGATACCTATACATCTCA 1249
 QY 386 AlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGln 402
 DB 1250 GCATG---GCCGAGACAGCTGCTGACCTCAGCGTCTTCAAGACCTGCAAGTAATCCGGG 1308
 QY 403 403
 DB 1309 ACGAATTCTGCACAATGAGCGCTACTGCTGACCTGCAGAGGCTGGGATCAGCTGCT 1368
 QY 409 409
 DB 414 414

DB 1369 GGGGCTGCTCACTGAGGAACCTGGGCACTGGAGCTGGCCCTCATCCACCATACACCA 1428
 QY 415 415
 DB 1429 CCTGCTGCTGGACACAGGCTGCCCTGGACACAGCTCTTTCGGAACCCGACCAAG--- 1485
 QY 432 ValAsnGlnProAspValArgProGln 1485
 DB 1486 ---TCTGCTCCACACCTCCCAACCGCCAGAGAGACAGATGTGTGGCGAGGCTGGCTG 1542
 QY 444 ProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlnAlaGlyLys 463
 DB 1543 CCA---CCAGCTGTGCCCGGAGGCGACTGCTGGGTCCAGGCGCCAC 1587
 QY 464 ThrLeu 473
 DB 1588 CCACTGTGTCACTGCAGCCAGTTCTTCCGGGGCCAGAGATGCTGTGAGGAATGCCAGT 1647
 QY 474 474
 DB 1648 ACTGACAGGCTGCCAGGAGATGTGAAATGCCAGGACACTGTTGCCGTGCCACCTGA 1707
 QY 483 ValGluAsnProGluTyrLeuThr 1767
 DB 1708 GTGTACAGCCCGAGATGCTCAGTACAGCTGTTTGGACCGGAGGCTGACAGCTGTGGC 1767
 QY 495 AlaAlaPro-GlnProHisPro 507
 DB 1768 CTGGGCCACTATAGAACCTCCCTTCTGCTGCGGCCCGCCGACGCGGTGAAAC 1827
 QY 507 AlaPheAspAsnLeuTyrTyrTrpAspGlnAspProGluArgGlyAla---ProPr 526
 DB 1828 TGACCTCTCTACATGCCCATGTGG---AAGTTCCAGATGAGAGGCGCCATGCCAGCC 1884
 QY 526 OserThrPhe 535
 DB 1885 TTGCCCATCAACTGCACCCACTCTGTGTGACCTGATGACAGAGGCTGCCCGCGCA 1944
 QY 535 AGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAlaIleArgSerProse 555
 DB 1945 GCAGAGAGCCAGCCCTCTG---ACGTCCCTGAGAGGACCGCCGCTGCCAG 1995
 QY 555 rProSerThrGlnProTryptLeuHisValAsnAlaIleGlnGluAlaArgArgLeuAs 575
 DB 1996 CCCAGACACAGCCCTGGGAGCATGTGAATGCCATCCAGAGGCGCCGCTCTCCAGAA 2055
 QY 575 nLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIleSerGluMetPh 595
 DB 2056 CTGAGTAGAGACACTGCTCTGATGATGAATGAACAGTGAAGTCACTTCAGAAATGTT 2115
 QY 595 eAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuTyrLysGlnLeuArg 615
 DB 2116 TGACCTCCAGAGCGGACCTGCTGCTGATGAATGAACAGTGAAGTCACTTCAGAAATGTT 2175
 QY 615 gGlySerLeuThrLysLeuGlyGlyProLeuThrMetMetAlaSerHisTyrLysGlnI 635
 DB 2176 GGGAGGCTCACCAGAGTCAAGGGCCCTTGACCATGATGTGCGCAGCCATCAACAGCA 2235
 QY 635 sCysProProThrProGluThrSerCysAlaIleGlnIleIleThrPheGluSerPhe 655
 DB 2236 CTGCCCTCCAAACCCGGAACCTCTGCTGCAACCCCAATATACCTTTAAGATTTCAA 2295
 QY 655 sGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTyrPgluProValGlnI 675
 DB 2296 AGAAGACTGAAGAGACTTGTGCTGTCAACCCCTTGAATGCTGCTGGAGGACAGTCCAGA 2355
 QY 675 uGlyAlaProProProAla 682
 DB 2356 GTGAGACCGGCCAGATGAGCT 2377
 RESULT 11
 US-08-229-515A-14

```

Sequence 14, Application US/08229515A
Patent No. 5518885

GENERAL INFORMATION:
APPLICANT: RAZIUDIN
APPLICANT: SARKAR, FAZLUL H
TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
TITLE OF INVENTION: NEOPLASTIC DISEASE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG PC
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia
COUNTRY: usa
ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/229,515A
FILING DATE: 19 APR 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: PERRYMAN, DAVID G
REGISTRATION NUMBER: 33,438
REFERENCE/DOCKET NUMBER: 1414, 608
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-688-0770
TELEFAX: 404-688-9880
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 3955 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-229-515A-14

Alignment Scores:
Pred. No.: 4,98e-129 Length: 3955
Score: 1968.00 Matches: 443
Percent Similarity: 35.87% Conservative: 24
Best Local Similarity: 34.02% Mismatches: 83
Query Match: 52.63% Indels: 752
DB: 1 Gaps: 4

US-09-821-883-2 (1-690) x US-08-229-515A-14 (1-3955)
QY 3 AlaAlaProLeuLeuLeuAlaAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 22
| | | | | : : : : : | | | | | : : : : : | |
DB 11 GCCGCAATGATCATCATGAGAGTGGCGGCTGGTGCCGCTGGGGTTCTCTCGCCCTC 70
DY 23 PhePheTrpPheLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
| | | | | : : : : : | | | | | : : : : : | |
DB 71 -----CTGCCCGCCCGGAAATGCGCGGCAC 94
QY 43 GluValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
| | | | | : : : : : | | | | | : : : : : | |
DB 95 CAAAGTGTATACCGGACACAGCATGAATGGCGGCTCCCTGCCATGCTCGAAGACCCACCTG 155
QY 63 AspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGluLeuThr 82
| | | | | : : : : : | | | | | : : : : : | |
DB 155 GACATGCTCCGCGACACTGTACACAGGGCTGTCAGGTAGTCAGAGGCAACTGTGAGCTTACC 214
QY 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGlyTyr 107
| | | | | : : : : : | | | | | : : : : : | |
DB 215 TACGTGCTTGCACAAATGCGAGGCTCTCATTTCTCTGCAGACATCCAGGAATTCAGAGTTAC 277
QY 103 ValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIleValArg 122
| | | | | : : : : : | | | | | : : : : : | |
DB 275 ATGCTCATTCGCTCACACACAGGTGAAGGCGCTCCACATCGCAAAAGCTGCGCATGTGTGAGA 330

```

Db 1415 AGTGGATTGGCTCTGATTCAACCGCCATCTCTGTTGTACACACTGACTTGG 1474
Oy 329 ----- 329
Db 1475 GACCAGCTTTCCGGAAACCAACATCAGGCCCTGCTCCACAGTGGGAACCGCCGGAAGAG 1534
Oy 329 ----- 329
Db 1535 GACTTGTGCGTCTCGACGGGCTTGTGTGTAACACTGTGTGCCACAGGGCACTGCTGG 1594
Oy 329 ----- 329
Db 1595 GGGCCAGGGCCCAACCCAGTGTCTCACTGCAATCTCTCGGGCCAGAGTGTGTG 1654
Oy 329 ----- 329
Db 1655 GAGAGTGCAGAGTATGGAAGGGGCTCCCGGAGTATGTGAGTGCACAAAGCCTGTCTG 1714
Oy 329 ----- 329
Db 1715 CCGTTCACCCCGAGTGTACGCTCAAAACAGCTCAGAGACTGCTTTGGATCGAGGCT 1774
Oy 329 ----- 329
Db 1775 GATCAGTGTGACGCTGCGCCCACTACAGAGACTCGTCTCTGTGTGCTCGCTGCCCC 1834
Oy 329 ----- 329
Db 1835 AGTGTGTGAACCGGACCTCTCTACATGCCATCTGGAAGTACCGGATGAGAGGGC 1894
Oy 329 ----- 329
Db 1895 ATATGCCAGCCGTGGCCCATCACTGCACCCACTCTGTGTGATGTGATGAAGAGGC 1954
Oy 329 ----- 329
Db 1955 TGCCACAGCAGACAGAGACCGCCGGTGACATTCATTCATTCGCACTGAGAGGGCTC 2014
Oy 329 ----- 329
Db 2015 CTGCTGTCTGATCTTATAGTGTGTGTGTGTAATCTTAACAACGAGACAGAGAAG 2074
Oy 329 ----- 329
Db 2075 ATCCGGAATATACGATGCGTAGGCTGTGAGAACTAGTTAGTAGAGCGCTGACG 2134
Oy 329 ----- 329
Db 2135 CCCAGCGGACCATGCCCAACAGGCTCAGATCGGATCTTAAAGAGACGAGCTAAGG 2194
Oy 329 ----- 329
Db 2195 AAGGTGAAGGTGTGATCAGAGGCTTTTGGCACTGTCAAGAGGCACTGTGGATCCA 2254
Oy 329 ----- 329
Db 2255 GATGGGAGATGTGAATAATCCCGTGGCTATCAAGGTGTGTAGAGAAACACATCTCT 2314
Oy 329 ----- 329
Db 2315 AAAGCCAACAAGAAATTTAGATGAAGCGTATGTGATGCTGTGGTGTCTCCGTAT 2374
Oy 329 ----- 329
Db 2375 GTGTCCCGCTCGTGGGATCTGGCTGACATCCAGATACAGCTGATGATGATG 2434
Oy 329 ----- 329
Db 2435 CCCTAAGGCTCCTTCTGACATGTCCGAGAACACCGAGGTGCTTACGCTCCAGAGAC 2494
Oy 329 ----- 329

Db 2495 CTGCTCAACTGTGTGTCTTCAAGATTGCCAAGGGATGAGTACCTGAGAGACTGCGGCTT 2554
Oy 329 ----- 329
Db 2555 GTACACAGGAGCACTGGCTGCCCGGAATGTGTCTAGTCAAGAGTCCCAACACATCAAGATT 2614
Oy 329 ----- 329
Db 2615 ACAGATTTCCGGGCTGCTCGGCTGTGTGACATTGATGAGACAGATACCATGAGATGGG 2674
Oy 329 ----- 329
Db 2675 GCGAAGGTGCCATCAAAATGATGAGATGGAATCTAATCTACAGACCGCGTTCACCAT 2734
Oy 329 ----- 329
Db 2735 CAGAGTATGTGTGAGCTATGAGTACATCTGTGTGAGAGCTGATGACTTTTGGGGCCAAA 2794
Oy 329 ----- 329
Db 2795 CATTACGATGGAATCCAGCCCGGAGATCCATGATTGCTGGAGAAGGAGAACGCTTA 2854
Oy 329 ----- 329
Db 2855 CCTCAGCTCCAAATCTGCACCAATTGATGTCTACATGATTATGCTCAAAATGTGATGATT 2914
Oy 329 ----- 329
Db 2915 GACTGTGAATGTGCGCCGAGATTCCGGAGTTGGTGTCAAAATTTTACGATATGGGAGG 2974
Oy 329 ----- 329
Db 2975 GACCCCGAGGCTTTTGTGTATCAAGAACGAGACTTGGGCCCATCCAGCCCATGAGAC 3034
Oy 329 ----- 329
Db 3035 AGTACTTCTACCGTTCACTGTGGAAGATGATGACATGGCTGACCTGATGACGCTGAA 3094
Oy 330 ----- 329
Db 3095 GAGTATGTGTGCCCCGAGGAGGATTTCTTCCCGGAGCCCTACCCAGGCACTGGGAGC 3154
Oy 334 ----- 329
Db 3155 ACAGCCCATGGAAGCGCCACCTCTCCACAGAGGTGAGGTGTGCTGACACTG 3214
Oy 354 ----- 329
Db 3215 GGCCTGAGCCCTCGGAGAAGGGCCCCCAGATCTCCTGCTCCCTCGGAAGGGCT 3274
Oy 374 ----- 329
Db 374 GlySerAspValPheAspGlyAspLeuGlyMetGlyValAlaLeuGlyLeuGlnSerLeu 383
Oy 3275 ----- 329
Db 3275 GGCCTCGATGTGTGTGATGTGACCTGCAATGAGGGGTAACAAAGGCTGCAAGACCTC 3334
Oy 394 ----- 329
Db 394 ProThrHisAspProSerProLeuGlnArgTyrSerGlnAspProThrValProLeuPro 413
Oy 3335 ----- 329
Db 3335 TCTCCACATGACCTCAGCCCTTCAACGGGTACAGCGAGGACCCACATTAATCTCTGCC 3394
Oy 414 ----- 329
Db 414 SerGlnThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlnTyrValAsn 433
Oy 3395 ----- 329
Db 3395 CCCGAGACTATGCTATGTGCTCCCTGCTGACAGCCCGCCAGCCGAGTATGTAAAC 3454
Oy 434 ----- 329
Db 434 GlnProAspValArgProGlnProProSerProArgGlnGlyProLeuProAlaAlaArg 453
Oy 3455 ----- 329
Db 3455 CAATCAGAGGTTCAGCCTCAGCTCTTAACCCAGAGGGGTCTGCTCTGTCGCG 3514
Oy 454 ----- 329
Db 454 ProAlaGlyAlaThrLeuGlnArgAlaLysThrLeuSerProGlyLysAsnGlyValVal 473
Oy 3515 ----- 329
Db 3515 CTTGCTGTGCTACTCTAGAAAGACCAAGACCTCTCTCTGGAAGAAATGGGGTGTG 3574
Oy 474 ----- 329
Db 474 LysAspValPheAlaPheAlaPheGlyAlaValGlnAspProGlnTyrLeuThrProGlnGly 493
Oy 3575 ----- 329
Db 3575 AAAGACGTTTTTCTTCCCTTCCGGGGGTGCTGTGAGAACCTGATTAATCTTACGAGAGAA 3634

Db 1055 TATGCTGGGCATGAGCACTTTCGAGGGCCAGGCCATCACCAGTACATGTCCAG 1114
OY 329 ----- 329
Db 1115 GAGTTGATGGCTGCAGAAAGATCTTTGGAGCCTGGCATTTTGGCGGAGAGCTTTGAT 1114
OY 329 ----- 329
Db 1175 GGGAGCCCTCTCCGGCATTTGCTCCGTGAGGCTGAGAGCTCAAGTGTTCGAAACC 1234
OY 329 ----- 329
Db 1235 CTGGAGAGATCACAGTTACCTGTACATCTCAGCATGGCCAGACACTCTCCGTGACCTC 1294
OY 329 ----- 329
Db 1295 AGTGTCTCCAGAACCTTCGAATCATTCGGGAGAGATTCACAGATGGCGGTACTCA 1354
OY 329 ----- 329
Db 1355 TTGACACTGCAGAGCCTGGGGATCCACTCGCTGGGCTGCGCTCAGTGGGAGCTGGGC 1414
OY 329 ----- 329
Db 1415 AGTGGATTGGCTGTGATTCAACCGCAAGCCCATCTGTCTTTGTACACTGTACCTTGG 1474
OY 329 ----- 329
Db 1475 GACCAGCTTTCCGGAACCCACATCAGGCCCTGTCCACAGTGGGAACCGGCCGGAAG 1534
OY 329 ----- 329
Db 1535 GACTTGTGCTCTCGAGAGGCTGTGTGTACTCACTGTGTGCCACAGGCACTGCTGG 1594
OY 329 ----- 329
Db 1595 GGGCAGGGGCCACCCAGTGTGTCACTGCAGTCACTTCTCTGGGGCCAGAGTGTGTG 1654
OY 329 ----- 329
Db 1655 GAGAGTCCGAGATGATGAAGGGCTCCCGGAGATGATGATGACAAAGCGTGTCTG 1714
OY 329 ----- 329
Db 1715 CCGTGTACCCCGAGTGTCAAGCTCAAAAGACTCAGAGACTGCTTTGGATCGAGGCT 1774
OY 329 ----- 329
Db 1775 GATGATGTGAGGCTGTGGCCACTACAGAGACTGTCTCTGTGTGTGGCTCGTGGCC 1834
OY 329 ----- 329
Db 1835 AGTGTGTGAACCGGACCTCTCTCATATGCCATCTGGAATGACCGGATGAGAGAGGC 1894
OY 329 ----- 329
Db 1895 ATATGCCAGCGGTGCCCATCAACTGCACCCACTCTGTGTGATCTGATGAACGAGGC 1954
OY 329 ----- 329
Db 1955 TGCCCAAGACAGAGAGAGCCGCCGTGACATTCATCTTGCAACTGTAGAGGGCTTC 2014
OY 329 ----- 329
Db 2015 CTGCTGTCTCTGATCTTATGTGTGTGTGTGATCTTATCAAAAGAGAGACAGAG 2074
OY 329 ----- 329
Db 2075 ATCCGAGATATACGATGCTAGGCTGTGACAGAAACTGATTAGTGAAGCGCTGACG 2134
OY 329 ----- 329
Db 2135 CCCAGGGAGCAATGCCCAACAGAGGCTCAGATGCGGATCTTAAAGAGAGAGGAGCTAAGG 2194

OY 329 ----- 329
Db 2195 AAGTGAAGGTCTTGGATCAGAGACTTTTGGCACTGTCTACAGGGCATCTGATCCCA 2254
OY 329 ----- 329
Db 2255 GATGGGAGAAATGAAATTCGCCGTGATCAAGGTGTGAGAGAAACACATCTCT 2314
OY 329 ----- 329
Db 2315 AAAGCCAAGAAATTTAGATGAACCGATGTGATGCTGTGTGGTTCCTGAT 2374
OY 329 ----- 329
Db 2375 GTGTCCGCTCTGGGCATCTGCTGATCCACAGTACAGCTGTGACACAGCTTATG 2434
OY 329 ----- 329
Db 2435 CCTACGGCTGCTTGTGACCATGTCCGAGAACCCAGAGTGGCTGAGCTCCAGGAC 2494
OY 329 ----- 329
Db 2495 CTGCTCAACTGTGTGTTCAGATTGCCAGGGGATGAGCTACTGAGAGCTGTGGCTT 2554
OY 329 ----- 329
Db 2555 GTACACAGGACCTGGCTGCCCGGATGTGTCTAGTCAAGATCCCAACAGTCAAGATT 2614
OY 329 ----- 329
Db 2615 ACAGATTTGGGGCTGGCTGGCTGCTGCTGACATGATGAGACAGAGTACATGAGATGG 2674
OY 329 ----- 329
Db 2675 GGCAGGTGCCCATCAATGATGATGATGATCTTCTCAGAGCCCGTCAACCAT 2734
OY 329 ----- 329
Db 2735 CAGAGTATGTGAGACTATGAGTACTGTGTGGAGCTGATGACTTTTGGGCCAAA 2794
OY 329 ----- 329
Db 2795 CCTTACGATGATCCAGCCCGGAGATCCCTGATTTCTGTGGAGAGCAAGCCCTA 2854
OY 329 ----- 329
Db 2855 CCTGAGCTCCAAATCTGCACACATGATGTCTACATGATTAATGTCAAATGTTGATGATT 2914
OY 329 ----- 329
Db 2915 GACTGTGAATGTGCCCCGAGATTCCGGAGTGTGTGAGAAATTTTCAGATGCGATGCGAGG 2974
OY 329 ----- 329
Db 2975 GACCCCGAGGTTTTGTGTATCCAGAACGAGACTTGGGCCCATCCAGCCCATGAGC 3034
OY 329 ----- 329
Db 3035 AGTACTTCTACCGTCTACTGTGAGATGATGATGATGGTGTGACTGTGATGACGCTGAA 3094
OY 330 ----- 330
Db 3095 GAGTATGTGTGCCCGCAGAGGATTTCTCCCGGACCTTACCCAGGACATCGGAGGAGC 3154
OY 334 MetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThrLeu 353
Db 3155 ACAGCCCATGAAAGGCCACCCAGCTGCTCCACAGAGAGTGGAGGTGTGCTGACACTG 3214
OY 354 GlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGlyAla 373
Db 3215 GGCTTGAAGCCTCTGGAGAGAGGGCCCCCAGATCTCTACTGCTCTCTGGAGAGGGGCT 3274

QY 374 GlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaValGlyLeuGlnSerLeu 393
DB 3275 GGCCTCCATGTTGTTGATGGTACCTGGCAATGGGGGTAAACCAAGGGCTGCACAGCCTC 3334
QY 394 ProThrHisAspProSerProLeuGlnArgTyrSerGlnAspProThrValProLeuPro 413
DB 3335 TCTCCACATGACCCACCCCTCTACAGCGGTACAGCAGACGCCACATTAATCTCTGCC 3394
QY 414 SerGlnThrAspGlyTyrValAlaProLeuThrGlySerProGlnProGlyTyrValAsn 433
DB 3395 CCCGAGACTGATGGCTATGTTGCTCCCTGCTGCAGCCCCCGAGCTATGTATGATAC 3454
QY 434 GlnProAspValArgProGlnProProSerProArgGlnGlyProLeuProAlaAlaArg 453
DB 3455 CAATCAGAGGTTCAGCTCAGCTCCTTTAACCCACAGAGGTCTCTGCTCCCTGCTCCG 3514
QY 454 ProAlaGlyAlaThrLeuGlnArgAlaValThrLeuSerProGlyTyrAsnGlyValVal 473
DB 3515 CCTGCTGCTGCTACTAGAAAGACCAAGACTCTCTCTGCGAAGAAATGGGGTTGTC 3574
QY 474 LysAspValPheAlaPheGlyGlyValAlaValGlnAspProGlyTyrLeuThrProGlnGly 493
DB 3575 AAAGACGTTTTTGGCTTGGGGGTGCTGTGGAGAACCTGATACCTTAAGTACGAGAA 3634
QY 494 GlyAlaAlaProGlnProGlnAspProProProAlaPheSerProAlaPheAspAsnLeuTyr 513
DB 3635 GGCACCTGCTGCTGCGCCACCTTCTGCTGCTTACAGCCAGCTTGTGACAACTCTAT 3654
QY 514 TyrTrpAspGlnAspProProProGlyAlaValProProSerThrPheLysGlyThrPro 533
DB 3695 TACGTGGACCAAGAACTATCGAGCAGGCGCTCCACCAAGTAACTTGAAGGAGACCC 3754
QY 534 ThrAlaGlnAspProGlyTyrLeuGlyLeuAspValProAlaAlaAlaProAlaValSer 553
DB 3755 ACTCAGAGAACCTGAGTACTAGGCTGATGATCTGATATAGACGCTGCGAGACGT 3814
QY 554 Pro-----SerProSerThrGlnProTrpGlu 562
DB 3815 CCTGTGCTTCAGAGTGGGGAAGGCTGACTGTGTGCTCATCGCCCAAGAGAGAG 3874
QY 563 HisVal 564
DB 3875 AGGGTC 3880

RESULT 13
US-08-422-108-2
Sequence 2, Application US/08422108
Patent No. 6015567
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,108
FILING DATE: 14-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA: 07/354319
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1872 nucleotides
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
US-08-422-108-2
Alignment Scores:
Pred. No.: 6.44e-107 Length: 1872
Score: 1647.00 Matches: 380
Percent Similarity: 57.68% Conservative: 22
Best Local Similarity: 54.52% Mismatches: 146
Query Match: 44.05% Indels: 150
DB: Gaps: 20
US-09-821-883-2 (1-690) x US-08-422-108-2 (1-1872)
QY 41 SerThrGlnValCysThrGlnGlyThrAspMetLysLeuArgLeuProAlaSerProGlnThr 60
DB 1 AGCACCCAGAGGUGUCACCGGACAGACAGACAGAGGUGGUGGUGGUGGUGGUGGUGGUGG 60
QY 61 HisLeuAspMetLeuAspGlnHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
DB 61 CACUGGAGACAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 120
QY 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100
DB 121 CUCACUACUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 180
QY 101 GlyTyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 181 GGCUGAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 240
QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 241 GUGGAGGAGACCCAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 300
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnLeuGln 160
DB 301 CCGCUGAACAAUACCCAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 360
QY 161 LeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
DB 361 CUUGAGAGCCUACAGAGAUUUGAAGAGAGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 420
QY 181 CysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnAlaLeu 200
DB 421 UGCUACAGGAGACAGAUUUUGGAGAGACAUUUCACAAAGAACACAGGUGGUGGUGGUGGUGG 480
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysLysGly 220
DB 481 ACACUGAUGAGACCAACCGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGGUGG 540
QY 221 SerArgCysTrpGlyGlySerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240
DB 541 UCCCGUGGUGGAGAGAGUUCUGAGAGAUUGUGAGACCGUAGCGCCACUUGUGGUGGUGGUGG 600
QY 241 GlyLysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260

LENGTH: 1872 nucleotides
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 US-08-422-734-2

Alignment Scores:
 Pred. No.: 6,44e-107 Length: 1872
 Score: 1647.00 Matches: 380
 Percent Similarity: 57.68% Conservativeness: 22
 Best Local Similarity: 54.52% Mismatches: 146
 Query Match: 44.05% Indels: 150
 DB: 4 Gaps: 20

US-09-821-883-2 (1-690) x US-08-422-734-2 (1-1872)

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
 Db 1 AGCACCCAGAGUGGACCGGACAGACAGAAAGCGCGCCGCCAGGCCAGCCAGACC 60
 Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValGlnGlnValGlnGln 80
 Db 61 CACCGGACAGUGCCUCCGCCACUCCUCCAGAGGCGUGCCAGUGUGUAGAGAAACCGGGA 120
 Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
 Db 121 CUCACCUACUCCGCCACCAAGUCCAGCCUCCUCCUCCAGAGAAUCCAGAGAGUGGCG 180
 Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
 Db 181 GCGUACGGGCGUACGCGCACACCAAGUGAGGCGUCCCGCCAGGCGUGCGGAGU 240
 Qy 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlnLys 140
 Db 241 GUGGACGACCCAGCCAGCUCUUGAGACAAUAGGCCCGGCCGUCUAGACAAUAGGAGAC 300
 Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
 Db 301 CCGCUGAACAAUACCAACCCUCCUCCAGAGGCGCCGCCAGAGGCGUGGAGGCGUGCAG 360
 Qy 161 LeuArgSerLeuThrGlnIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180
 Db 361 CUCCAGAACCCUCCAGAGUCCUAGAGAGGCGUCCUCCAGGAGAACCCAGCCAGCUC 420
 Qy 181 CysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
 Db 421 UCGUACAGAGACAGAUUUGUGAGAGACAAUCCUCCAGAACACACACAGCUGGCGUC 480
 Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
 Db 481 ACACUGAUAAGACACCAACCGCUCUGGCGCCAGCCCGUCCGUAUUGUUAAGGCG 540
 Qy 221 SerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThrValLysAla 240
 Db 541 UCCGCCUCCUGGAGAGAGUCCUAGAGUCCUAGAGCCUCCAGCCAGCCAGCUCUCCGUGCC 600
 Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
 Db 601 GUGGCGUCUGCCCGGUGGAGAGGCGCCAGCUGCCAGCUGGCGUCCAGUAGAGUGGU 660
 Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
 Db 661 GCGGCGUCACAGGCGCCCAAGACACUCCUCCGCGCCUCCGCGUCCAGUUAACCAAGU 720
 Qy 281 GlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlnSer 300
 Db 721 GGCACUCUGAGACUCCAGCCAGCCUCCUCCUCCAGCUCUACACAGACAGCGUUAAGUCC 780
 Qy 301 MetProAsnProGlnGlyArgTyrThrPheGlnAlaSerCysValThrAlaLysProTyr 320
 Db 781 AUGCCCAUCCGAGAGGCGGUAUACAUUCCGCGCCAGCUGUGUACUCCGUGUCCUAC 840
 Qy 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArg 340

Db 841 AACUACCUUUCAGGAGGUGGAGUCCUGACAC-----CUCGUCUGCCCGCCAGACAAAC 894
 Qy 341 SerSerSerThrArgSerGlyGlyLysP----- 350
 Db 895 CAAGAGGUGACAGCAGAGAGUAGAACACAGCGGUGUGAGAGUCCAGACAGCCUGGCC 954
 Qy 351 -----LeuThrLeuGlyLeuGlnProSerGlnGlnValAlaProArgSerProLeu 367
 Db 955 CGAGUGUCUAGUGUCUGGCGACAGACUCCUCCAGAGAGUCCAGGAGGAGUCCAGUCCAGUCC 1014
 Qy 368 AlaProSerGlnGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAla 387
 Db 1015 AAUACCCAGAGUCCUGGCGUCCAGAAAGAUCCUUGGAGCCUGGACAUUU----- 1065
 Qy 388 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerLysAsp 407
 Db 1065 ----- 1065
 Qy 408 ProThrValProLeuProSerGlnThrAspGlyTyrValAlaProLeuThrCysSerPro 427
 Db 1066 -----CUGCCGAGAGCUCUUGAGGAGACCCAGCCUCCACACUCCCGCCGUC 1113
 Qy 428 GlnProGln-----TyrValAsnGln 434
 Db 1114 CAGCCAGAGACACUCCAGAGUCCUUGAGACUCCUGAAGAGAUACAGAGUACCUAUUA-CAU 1172
 Qy 435 ProAspValArgProGlnProProSerProAlaGlyGlyProLeuProAlaAlaArgPro 454
 Db 1173 CUCACAGGCGCGGACACUCCUGGACUCCUCCUCCUCCGAAAGCUGCAAGUUAUCCG 1232
 Qy 455 AlaGlyAlaThrLeuGlnArgAlaLysThrLeuSerProGlyLysAsnGlyValValLys 474
 Db 1233 GGGAGAAUUCGACAAUAGGCGCCUACUCCGUGACCCUCCAGAGGCGUGS----- 1283
 Qy 475 AspValPheAlaPheGlyGlyAlaValGlnAsnProGluTyrLeuThrProGlnGlyGly 494
 Db 1284 -----CAUACAGCGUGGUGGCGUGCG-----CUCACUGAGGAGAAUCCUGG 1322
 Qy 495 AlaAlaProGlnProHisProPro----- 502
 Db 1323 CAGUGAGCUGGCCUCCUACUCCAGCAUACACACCCAGCCUCCUCCUCCAGACAGGUGCCUG 1382
 Qy 503 ---ProAlaPheSer---ProAlaPheAspAsnLeuTyrTyrTrpAspLysProPro 520
 Db 1383 GGACAGCUCUUCGGAACCGCCAGCAGCAGCUCUGUCCACACUAGGCCAA-----CCG 1433
 Qy 521 GlnArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGlnAsnProGlu--- 539
 Db 1434 GCCAGAGGACAGUCCUGGCGGAGGCGCCUCCGCGCCAGCAGCUGUGCGCCGAGAGGCA 1493
 Qy 540 TyrLeuGly-----LeuAspValProAlaAlaAlaPro 550
 Db 1494 CUGCGGGGUGCCAGGCGCCAGCAGUCCUCCAGCAGCCAGUCCUCCGCGGCCAGGA 1553
 Qy 551 AlaArgSerProSerProSerThrGln-----ProTrpGlnHisValAsnAlaIleGln 568
 Db 1554 GUGCGUGGAGAAUCCCGAGUCCUCCAGGCGUCCCGAGGAGUUAUUGAAGGCCAGGA 1613
 Qy 568 nGluAlaArgArgLeuLeuAsnLeuSerArgAspThrAlaAlaGlnMetAsnGlnThrVal 588
 Db 1614 C----- 1614
 Qy 588 GlnValIleSerGlnMetPheAspLeuGlnLupProThrCysLeuGlnThrArgLeuGln 608
 Db 1615 -----UGUUUGCGCGUCCAGCCUCCUA 1634
 Qy 608 uLeuTyrLysGlnGlyLeuArgGlySerLeuThrLysLeuLysGlyPro-----Le 625
 Db 1635 GUGUACGCCCGAG-----AANCGCUCAGUAGACUCCUUUU---GGACCGGAGCGUAGACA 1685
 Qy 625 uThrMetAlaSerHisTyrLysGlnHisCysAspProThrProGluThrSerCysAl 645

Db 1686 GUGUGGCCUGGCCACUUAAGAC-----CCUCCUUCUGCGGCCGCCGCC 1739
 QY 645 arthGluIleIleThrPheGluSerPheLysGluAsnLeuLysAspPheLeuVal11 665
 Db 1740 CAGC-----GUGUGAACCUGACCCUCCUCCAU 1769
 QY 665 eProPheAspCysTrpGlu---ProValGlnGluGlyAlaProProPro 680
 Db 1770 GCCCAUC-----UGAAGUUCUCCAGUAGAGGCGCGCAUGCCAGCCU 1812
 RESULT 15
 US-09-146-283-1
 ; Sequence 1, Application US/09146283
 ; Patent No. 5976546
 ; GENERAL INFORMATION:
 ; APPLICANT: Laus, Reiner
 ; APPLICANT: Ruegg, Curtis L.
 ; APPLICANT: Wu, Hongyu
 ; TITLE OF INVENTION: Immunostimulatory Compositions
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Dehlinger & Associates
 ; STREET: 350 Cambridge Ave. Suite 250
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/146,283
 ; FILING DATE: 03-SEPT-1998
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Judge, Linda R.
 ; REGISTRATION NUMBER: 42,702
 ; REFERENCE/DOCKET NUMBER: 7636-0010.21
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-324-0880
 ; TELEFAX: 650-324-0960
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1588 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHEICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF
 ; IMMEDIATE SOURCE:
 ; LIBRARY: prostate carcinoma LncAP.FGC; PBMC
 ; US-09-146-283-1
 Alignment Scores:
 Pred. No.: 2.05e-45 Length: 1588
 Score: 770.00 Matches: 257
 Percent Similarity: 42.14% Conservative: 54
 Best Local Similarity: 34.82% Mismatches: 159
 Query Match: 20.59% Indels: 268
 DB: 2 Gaps: 28
 US-09-821-883-2 (1-690) x US-09-146-283-1 (1-1588)
 QY 1 MetArgLaaIaPLeuLeuLeuLaaArgLaaLaaSerLeuSerLeuGlyPheLeuPhe 20
 Db 15 ATGAGAGCTGACCCCTCTCTGCGCAGGCGAAGACCTTACCTTGGCTTCTTT 74

QY 21 LeuLeuPheThrPLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
 Db 75 CTGCTTTTCTGCTAGACCGAAGAGTACTACCAAGAGATTG----- 119
 QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProLaaSerProGluThr 60
 Db 119 ----- 119
 QY 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
 Db 119 ----- 119
 QY 81 LeuThrTyrlleuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGln 100
 Db 119 ----- 119
 QY 101 GlyTyrlValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
 Db 119 ----- 119
 QY 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrlAlaLeuAlaValLeuAspAsnGlyAsp 140
 Db 120 -----AAGTTGTGACTTTGGTGTTCGCGATGAGAC 152
 QY 141 -----ProLeuAsnAsn-----ThrThrProValThrGlyAlaSer---ProGlyGly 155
 Db 153 CGAAGTCCCATTTGACACCTTTCCCATGACCCCATTAAGAAAGATCCATGCGCCACAAGA 212
 QY 156 LeuArgGluLeuGlnLeuAsnArgSerLeuThrGlnLeuLysGlyValLeuIleGln 175
 Db 213 TTTGGCCAACTCCACCCAGCTGGGCATGAGCATTAATGAACTTGAGACTATATAGA 272
 QY 176 ArgAsnProGlnLeuCysTyrlGlnAspThrIleLeuThrLysAspIlePheHisLysAsn 195
 Db 273 AAGAGATATAGAAATTCCTGAATGAGTCC-----TATTAACAT 311
 QY 196 AsnGlnLeuAlaLeuThrIleAspThrAsnArgSer---ArgAlaCysHis-----P 213
 Db 312 GAAACAGTTATATATTCGACACACAGACGTGACCGGACTTGATGAGCTATATACAAC 371
 QY 213 rocYSerPrometCysLysGlySerArgCysTrpGlyLysSerSerGlyLysAspCysGlns 233
 Db 372 CTGGCAGCCCTGTT-----TCCCCAGAAGGTGTACG 404
 QY 233 eIleuThrArgThrValCysAlaGlyLysAlaArgCysLysGlyProLeuProThrA 253
 Db 405 ATCTGAA-----TCCATATCTACTC 425
 QY 253 spCysCysHisGlnGlnCysAlaAlaGlyCys-ThrGlyProLysHisSerAspCysLeu 272
 Db 426 TGGCAGCCCAT-----CCCGGTGCACACAGTCTCTTCTGAAAGACAGCTTG 473
 QY 273 AlaCysLeuHisPheAsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThr 292
 Db 474 CTATACCTGCTTTCAGC-----AACTGCCCTCGTTTCAAGAA 512
 QY 293 TyrAsnThrAspThrPheGluSerMetProAsnProGluArgTyrlThrPheGlyAla 312
 Db 513 CTTCAGAGTGAAGATTGAATCAGAGAAATCCAGAAAGAGCTGCAC----- 560
 QY 313 SerCysValThrAlaCysProTyrlAsnTyrlLeuSerThrAspValGlySerGlyAlaGly 332
 Db 561 -----CCTTAAGAGATTTATACCTACCTGGGAAAACTTTCAGA 602
 QY 333 GlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThr 352
 Db 603 -----TTACAT-----GCCAGAGACTTTT 623
 QY 353 LeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGly 372
 Db 624 -----GGA 626

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 8, 2003, 12:30:24 ; Search time 2116 Seconds

(without alignments)
5281.137 million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARASLSLGLF.....EPVOEGAPPPRAAHNNHHH 690

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-xljp
-O=/cgn2_1/USPRO.spool/US09821883/unat_28042003.104749.14578/app-query.fasta.1.839
-DB-EST -QFMT-fastcap -SUFFI-p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-Biosum62 -TRANS-human40.cdi -LIST=45
-DOCALLIGN=200 -SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT-pct -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN.1.1.2814 @unat.28042003.104749.14578 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGESOURCY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
-WARN_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_estl: *
10: gb_estl2: *
11: gb_hic: *
12: gb_estl3: *
13: gb_estl4: *
14: gb_estl5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1288	34.4	1016	13	BM562913
2	1176.5	31.5	1129	14	BM802792
3	1163	31.1	4715	11	AF318349
4	1087	29.1	653	14	BM721340
5	1046	28.0	569	9	AL701765
6	976	26.1	943	14	BO958632
7	924	24.7	813	13	BI557797
8	915	24.5	614	10	BM370693
9	912	24.4	791	13	BI154872
10	901	24.1	491	14	BM790293
11	889.5	23.8	698	13	BI649877
12	881	23.6	685	9	AU123871
13	868	23.2	609	10	AW701942
14	864.5	23.1	621	12	BC283493
15	853	22.8	887	14	BQ717097
16	848	22.7	786	13	BI155788
17	835.5	22.3	649	9	AA496412
18	817	21.9	552	12	BG277542
19	811	21.7	670	9	A1906012
20	808	21.6	876	14	BQ769889
21	802	21.4	514	12	BF998814
22	792.5	21.2	678	13	BI555157
23	788.5	21.1	499	12	BF746234
24	788.5	21.1	499	12	BF746304
25	781.5	20.9	657	10	BB620332
26	778.5	20.8	632	10	BB655484
27	774.5	20.7	767	12	BF964111
28	774	20.7	474	9	AL701569
29	747.5	20.0	466	9	A1906364
30	747	20.0	462	12	BF834727
31	735	19.7	956	9	AL552065
32	729	19.5	508	10	BE065758
33	722.5	19.3	923	14	BO888953
34	720	19.3	456	9	A1906050
35	719.5	19.2	2450	11	AK004911
36	719.5	19.2	2662	11	AK004883
37	719.5	19.2	2936	11	AK004944
38	717	19.2	537	14	BM793744
39	697	18.6	457	10	AM372023
40	690	18.5	729	12	BG174714
41	675	18.1	588	10	AM207707
42	675	18.1	660	10	BE218982
43	675	18.1	666	9	A1912784
44	675	18.1	672	10	BE671554
45	675	18.1	695	10	BE669962

ALIGNMENTS

RESULT 1
LOCUS BM562913 1016 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6566679 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5736771
ACCESSION BM562913
VERSION BM562913.1 GI:18809393
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1016)
NIH-MGC http://mgc.ncl.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12745 row: n column: 04
High quality sequence stop: 637.
Location/Qualifiers

FEATURES
SOURCE
1. 1016
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5736771"
/clone_1lb="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: PCMV-SPORT6;
Site.1: NotI; Site.2: SalI; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 218 a 322 c 278 g 198 t

ORIGIN

Alignment Scores:
Pred. No.: 3 41e-92 Length: 1016
Score: 1288.00 Matches: 253
Percent Similarity: 84.03% Conservative: 10
Best Local Similarity: 80.83% Mismatches: 23
Query Match: 34.45% Indels: 28
DB: 13 Gaps: 6

US-09-821-883-2 (1-690) x BM562913 (1-1016)

QY 93 LeuGlnAspIleGlnGluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGln 112
|||||
Db 72 TTACAGCATATCCAGAGGTGTCAGGCGTCACTGCTCAACCAAGTAGGCGAG 131
|||||
QY 113 ValProLeuGlnArgGluArgIleValArgGlyThrGlnLeuPheGlnAspAsnTyrAla 132
|||||
Db 132 GTCCCACTGCGAGAGGTGCGGATGTCGAGGACCCAGCTCTTGAGACAACCTATGCGC 191
|||||
QY 133 LeuAlaValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSer 152
|||||
Db 192 CTGCGCGGTGAGACATGAGACCCGCTGAAACAATACACACCCTGTACAGGGGCTCC 251
|||||
QY 153 ProGlyGlyLeuArgGluLeuGlnLeuArgSerLeuThrGluLeuLysGlyVal 172
|||||
Db 252 CCAGAGAGCGCTGCGGAGCTGACCTCGAAGCCTCAGAGATCTGAAGAGGGGCTC 311
|||||
QY 173 LeuIleGlnArgAsnProGlnLeuGlyTyrGlnAspThrIleLeuTyrPlyAspIlePhe 192
|||||
Db 312 TTGATCCAGCGGAACCCCGCTGCTGACAGGACGATTTGTGGAAGACATCTTC 371
|||||
QY 193 HisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHis 212
|||||
Db 372 CACAAGAACAACCAAGCGGCTCTCACACTGATACACACCAACCGCTCTCGGCGCTCCAC 431
|||||
QY 213 ProCysSerProMetCysLysGlySerArgCysTrrpGlyGlySerGlyAspCysGln 232
|||||
Db 432 CCCGTGTTCTCGATGTATAGAGGCTCCCGCTGCGGAGAGAGATCTCGAGGATTTGAC 491
|||||
QY 233 SerLeuThrArgThrValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThr 252
|||||
Db 492 AGCTTGAGGCGCACTGTCTGTGGCGGTGCTGCTCCCGCTCAAGGGGCCACTGCCACT 551
|||||
QY 253 AspCysCysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeu 272
|||||
Db 552 GACTGCTGCCATGAGCAGTGTGCTGCGGCTGACAGGGGCCCAAGCACTGTGACTGCTG 611
|||||

QY 273 AlaCysLeuHisPheAsnHisSerGlyTleCysGluLeuHisCysProAlaLeuValThr 292
|||||
Db 612 GCGTGCCTCCACTTCACACACAGGAGCATCTGTAGGTCGACATGCCAGCCCTGGTACC 671
|||||
QY 293 TyrAsnThrAspThrPheGlnSerMetCysProAsnProGluGlyArgTyrThrPheGlyAla 312
|||||
Db 672 TACAACACAGACAGGTTGATGCTCATGCTCCCAATCCCGAGGCGCGTATACATTTGGCGCC 731
|||||
QY 313 SerCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySerGlyAlaGly 332
|||||
Db 732 AGCTGTGAGCTGCTGCTGCTGCTTACAC-TACCTTTCTACGACGCGGATCTGCAC- 787
|||||
QY 333 GlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyAspLeuThr 352
|||||
Db 788 -CTGCTGCTCCCGCTCCACACCA-----GAGGTGACC 820
|||||
QY 353 Leu-GlyLeuGluProSerGlu---GluGluAlaProArgSerProLeuAlaProSerGlu 371
|||||
Db 821 GCAGACAGTGCAGACACACCGCTGTGAAAAAGTGCACACACCCCTGTGCCCA----- 875
|||||
QY 371 uGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu- 390
|||||
Db 876 -----GATGGCTAATGGGCTGGGCGCATGGAACACTTGGC 913
|||||
QY 391 -----GlnSerLeuProThrHisAspPro 398
|||||
Db 914 AAAAAGTGAAGGCGCAATTACCC---CATGCCCA 947
|||||

RESULT 2
BM802792 1129 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOUR_6459871 NIH_MGC_92 Homo sapiens CDNA clone IMAGE:5575087
DEFINITION 5', mRNA sequence.
ACCESSION BM802792
VERSION BM802792.1 GI:19119615
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1129)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12324 row: m column: 08
High quality sequence stop: 642.
Location/Qualifiers

FEATURES
SOURCE
1. 1129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5575087"
/clone_1lb="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: PCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 243 a 367 c 325 g 194 t

ORIGIN

Alignment Scores:

Oy 458 ThrleugluatgalaalysThleuSerProglytlysanglyvallyaspvalphe 477
 db 3781 ACTGGGAAGGCCCACTCTCCAGGAGGAATGGGCTGCTCAAGACCTTTT 3840
 Oy 478 AlaphesglylalaValaGluaspProglytThleuThProGlnGlyAlaAlaPro 497
 db 3841 GCCTTGGGGGGTCCCTGGAGAACCCGAGTACTGACACCCAGGAGAGCTGC-CCT 3899
 Oy 498 GlnProHlaProProProAlaPheSerProAlaPheaspasnleuTyTrpaspGln 517
 db 3900 CAGCCCAACCTCTCTGCTGCTGAGCCAGCTTCGACAACTTATTACTGGAGCAG 3959
 Oy 518 AspproProGluArglyAlaProProSerThrPheGlyGlyThProThAlaGluasn 537
 db 3960 GACCACACGAGCGGGGGCTCCAGCCAGCACTTCAAGGACACCTACGCGAGAGAAC 4019
 Oy 538 ProGluTyrluGlyLeuaspValProAlaAlaAlaProAlaArgSerProSerPro 556
 db 4020 CCAGAGTACTGGCTCTGAGCTGCCA-----GTGTACACCAAGAGCCAACTCCG 4070
 RESULT 4
 LOCUS BM721340 653 bp mRNA linear EST 01-MAR-2002
 DEFINITION UI-E-EO1-81b-p-15-0-UI r1 UI-E-EO1 Homo sapiens cDNA clone
 UI-E-EO1-81b-p-15-0-UI 5', mRNA sequence.
 ACCESSION BM721340
 VERSION BM721340.1 GI:19041189
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 653)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mssoares@blue.weeg.uiowa.edu
 Tissue Procurement: Dr. Gregg Hegeman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1. 653
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-E-EO1-81b-p-15-0-UI"
 /clone_1lb="UI-E-EO1"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)."
 /note="Organ: eye; Vector: pTR73-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-EO1 is a normalized cDNA library containing the
 following tissue(s): fetal eye. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pTR73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the

(dr)18 tail. The sequence tag for this library is
 CGCGTATACC. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI).
 BASE COUNT 130 a 226 c 188 g 108 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,78e-76 Length: 653
 Score: 1087.00 Matches: 202
 Percent Similarity: 96.19% Conservative: 0
 Best Local Similarity: 96.19% Mismatches: 6
 Query Match: 29.07% Indels: 2
 DB: 14 Gaps: 1
 US-09-821-883-2 (1-690) x BM721340 (1-653)
 Oy 347 GlyGlyGlyaspLeuThleuGlyLeuGluProSerGluGluAlaProArgSerPro 366
 db 3 GCGGGTGGGACCTGACACTAGGGCTGGAGCCCTTGAAAGAGAGGCCCCAGGCTTCCA 62
 Oy 367 LeuAlaProSerGluGlyAlaGlySeraspValPheaspGlyaspLeuGlyMetGlyAla 386
 db 63 CTGGCACCTCCGAAAGGGGCTGGCTCCGATATTGATGCTGACCTGGAGATGGGCA 122
 Oy 387 AlaGlyGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTySerGlu 406
 db 123 GCAAGGGGCTGCAAAAGCTCCGCCACACATGACCCACCCCTCTACAGCGGTACAGTGA 182
 Oy 407 AspproThrValProLeuProSerGluThrAspGlyTyValAlaProLeuThrCysSer 426
 db 183 GACCCCAAGTACCCCTGCTGCTGAGACTATGCTACGTTGGCCCCCTGACTGACG 242
 Oy 427 ProGlnProGluTyValaGlnProaspValArgProGlnProProSerProArgGlu 446
 db 243 CCCAGCCTGATATGTGAACACGACGATTTGGCCCCCAGCCCTTCCGCCCCGAGAG 302
 Oy 447 GlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaGlyThleuSer 466
 db 303 GGCCTCTGCTGCTGCTGCCGACCTGCTGCTGCTGCAAGAGCCCAAGACTCTCTCC 362
 Oy 467 ProGlytlysanglyvallyaspvalphealaphesglylalaValaGluasnPro 486
 db 363 CCAGGGAAGATGGGGTGTAAAGACTTTTGGGGGTGGCTGCAAGAACCC 422
 Oy 487 GluTyrluThrProGlnGlyAlaAlaProGlnProHisProProAlaPheSer 506
 db 423 GAGTACTGACACCCAGGAGAGAGCTGCCCTCAGCCCACTCTCTGCTTACAG 482
 Oy 507 ProAlaPheaspasnleuTyTrpaspGlnaspProProGluArglyAlaProPro 526
 db 483 CCAGCCTTCGACAACTCTATTACTGGACAGACCCACAGAGCGGGGGCTCCACCC 542
 Oy 527 SerThrPheGlyGlyThrProThAlaGluasnProGluTyrluGlyLeuaspValPro 546
 db 543 AGCACTTCAANAGGAGACCTTACGCGAGAGAACCCAGAGTACTGGCTGAGCTGCCA 602
 Oy 547 AlaAlaAlaProAlaArgSerProSerPro 556
 db 603 -----GTGTACACCAAGAGCCCAAGTCCG 626
 RESULT 5
 LOCUS AL701765 569 bp mRNA linear EST 22-MAR-2002
 DEFINITION DKEZ0686P06148_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DKEZ0686P06148 5', mRNA sequence.
 ACCESSION AL701765
 VERSION AL701765.1 GI:19685121
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 569)
 Duesterhoeft, A., Lauber, J., Mewes, H. W., Gassenhuber, J. and Wiemann
 S.
 TITLE EST (Duesterhoeft, et al.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Duesterhoeft A
 MFS

Am Klopferspitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No sl sequence available.
 This clone (DKFZp686P06148) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 Location/Qualifiers
 1..569

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZp686P06148"
 /clone_11b="686 (synonym: h1cc3)"
 /tissue_type="human skeletal muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector: pR1plex2; site_1: SfilI; site_2: SfilI;
 cDNA-collection"

BASE COUNT 115 a 187 c 156 g 110 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 2.7e-73 Length: 569
 Score: 1046.00 Matches: 188
 Percent Similarity: 98.95% Conservative: 0
 Best Local Similarity: 98.95% Mismatches: 2
 Query Match: 27.98% Indels: 1
 DB: 9 Gaps: 0

US-09-821-883-2 (1-690) x AL701765 (1-569)

QY 127 PHEGLUASPANTYTALEUAALAVALLUASPAANGLYASPPROLEUNASANTHTR 146
 DB 1 TTTGAGGACAC-TATGCGCTGGCGCTGAGACAATGGAGACCCCTGAAACAATACAC 59
 QY 147 PVOVALTHGLYALASERPROGLYLEUARGULUENLEUARGSERLEUHRGU 166
 DB 60 CCTGTACAGGGGCGCTCCCAAGAGGCGTGGAGAGCTGACCTGGAAGCCTCACAGAG 119
 QY 167 ILEULYSGLYVALLEULIEGLNARGASNPORGANLEUCYSTYGLINASPTHRI 186
 DB 120 ATCTGAAAGAGGGGCTGTGATCCAGCGGAACCCCACTGCTACACAGACAGATT 179
 QY 187 LEUTIRPLYSAPRIEPHEHISLYSASANSNGINLEUALALEUTHREULEASPTHAS 206
 DB 180 TTGTGGAAAGACATCTTCACAAAGAACACAGCTGGCTCTCACACTGATTAACACCAAC 239
 QY 207 ARGSEARGLACYSANISPROCYSSERPROMETCYSLYSGLYSERARGYSTRRPLGYU 226
 DB 240 CGCTCTCGGGCGCTGACCCCTGTTCTCCGATGTGAAGGGCTCCCGCTGGGGAGAG 299
 QY 227 SETSERGLUASPCYSGINSELEUTHRATGTRVALCYSAIAGLYCYSAIARQYS 246
 DB 300 AGTTGTAGAGATTGTCAAGAGCTGACGCCACTGTCTGTGCGGGTGGCTGCCGCCGCTGC 359
 QY 247 LYSGLYPROLEUPROTHRASPSCYSHTSGUINCYSALAIAGLYCYSTHRLGYPRO 266
 DB 360 AAGGGGCACTGCCCACTGACTGTGCTGAGACAGTGTGCTGCGGGCTGACGGGGCC 419
 QY 267 LYSHTSSERASPCYSLEUALACYSLEUHTSHPHASNHISSEGLYTLIEGYSLEUHTS 286
 DB 420 AAGCACTGTGACTGCTGGCTGCCCTCCCACTCAACACAGTGGAGTGTGAGCTGAC 479

QY 287 CYSPROALALEUALTHRYTASANTHRASPTHREGLUSERMETPROASNPGLUGLY 306
 DB 480 TGCCACAGCCCTGTCACCTACACAGACAGAGTTGATGTCATGCCAATGCCAGGAGC 539
 QY 307 ARGTYRTHREPHGLYALASERCYSVALTHR 316
 DB 540 CGGTATACATTGCGCGCAGCTGTGACT 569

RESULT 6

BO958632

LOCUS BO958632 943 bp mRNA linear EST 21-AUG-2002
 DEFINITION AGNC000RT.10061234 NCI_CGAP_Mam2 Mus musculus cDNA clone
 IMAGE:6485621 5', mRNA sequence.

ACCESSION BO958632

VERSION BO958632.1 GI:22374110

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov

Tissue procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LMNL at:

http://image.llnl.gov

Plate: LLM14027 row: 1 column: 06

High quality sequence stop: 579.

FEATURES

Location/Qualifiers

1..943

/organism="Mus musculus"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:6485621"

/clone_11b="NCI_CGAP_Mam2"

/tissue_type="tumor, biopsy sample"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 203 a 312 c 249 g 179 t

ORIGIN

US-09-821-883-2 (1-690) x BO958632 (1-943)

Alignment Scores:

Pred. No.: 1.85e-67 Length: 943
 Score: 976.00 Matches: 187
 Percent Similarity: 77.04% Conservative: 11
 Best Local Similarity: 72.76% Mismatches: 43
 Query Match: 26.10% Indels: 16
 DB: 14 Gaps: 2

QY 322 TYRLEUSERTHASPYVALGYSERGLYLAGLYGMYMETVALHISARHISARSR 341
 DB 18 TTCTTCTCCCAAGACCTGCGCTAGATGCTGGAGACACAGCCACGACAGACCGCAGC 77
 QY 342 SETSERTHRARGSERGLYGLYASPLEUTHRLEUGLYLEUGLUPROSERGLUGLU 361
 DB 78 TCGTGGGCGAGAGAGGGGGTGTGAGCTGACACACTGGGGCTGGAGCCCTCCGAAAGAG 137
 QY 362 ALAPROARGSERPROLEUALAPOSERGLUGLYLAGLYSERASPYVALPHEASPYL 381
 DB 138 CCCCACAGATCTCCACTGCTGCCCTCCGAAAGGGGCTGCTCGAGTGTGATGTGATG 197

VERSION AM370693.1 GI:6875347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS 1 (bases 1 to 614)
TITLE HCGP http://www.ludwig.org.br/ORESTES.
JOURNAL The FAPESP/LICR Human Cancer Genome Project
COMMENT Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL.
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1st2-QV1-BT0260-
011199-024-all&t3=1999-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 12
High quality sequence stop: 612.
Location/Qualifiers
1..614
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="BT0260"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site.1: Sma1; Site.2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 119 a 198 c 172 g 124 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7.5e-63 Length: 614
Score: 915.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.47% Indels: 0
Gaps: 0
US-09-821-883-2 (1-690) x AM370693 (1-614)
Qy 172 ValLeuIleGlnAArgAsnProGlnLeuGlyGlnAspThrIleLeuTrpLysAspIle 191
Db 17 GTCTTGATCCAGCGGACCCACGCTCTGCTACAGGACGACGATTGTTGGAAAGACATC 76
Qy 192 PheHisLysAsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerArgLacys 211
Db 77 TTCCCAAGAACACACACACCTGCTCTCACACTGATAGACACCAACCCGCTCGGGCTTGC 136
Qy 212 HisProGlySerProMetCysLysGlySerArgCysTrpGlyGlnLysSerGlnLysPcys 231
Db 137 CACCCCTGTTCCGATGCTGAAGGGCTCCGCTGCTGGGAGAGATTCTGAGAGATTGT 196
Qy 232 GlnSerLeuThrArgThrValCysAlaLysGlyCysAlaAlaArgCysLysGlyProLeuPro 251
Db 197 CAGACCCGACGCGCAGCTGTGTGTCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
Qy 252 ThrAspCysCysHisGlnGlnCysAlaAlaLysGlyCysThrGlyProLysHisSerAspCys 271
Db 257 ACTGACTCTGCTGCATGACAGAGTGTCTGCGGCTGCACGGGCGCCAAACACTGTACTGC 316
Qy 272 LeuAlaCysLeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuVal 291
|||||

Db 317 CTGGCGCTCTCCACTTCACACAGTCGATCTGTGAGCTGCACATGCCACCCCTGGTC 376
Qy 292 ThrTyrAsnThrAspThrPheGlnuSerMetProAsnProGlnuGlyArgTyrThrPheGly 311
Db 377 ACCATACAACAGACACCTTTGAGTTCATGCCATCCCGGGCCGGTAAATTCGCGC 436
Qy 312 AlaserCysValThrAlaCysProTyrAsnTyrLeuSerThrAspValGlySer 329
Db 437 GCCAGCTGTGACTGCTGCTGCTCCTACACTACTCTTCTACGGACGTGGATCC 490
RESULT 9
B1154872 791 bp mRNA linear EST 05-JUN-2001
LOCUS 602902857F1 NCL_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032238 5',
DEFINITION mRNA sequence.
ACCESSION B1154872 GI:14614873
VERSION B1154872
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
1 (bases 1 to 791)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgapbs-remail.nih.gov
Tissue Procurement: Lotzlar Hennighausen Ph.D., Chu Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LMNL1090 row: b column: 15
High quality sequence stop: 764.
Location/Qualifiers
1..791
/organism="Mus musculus"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone_id="IMAGE:5032238"
/clone_id="NCL_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lotzlar Hennighausen/Chu Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 159 a 264 c 215 g 153 t
ORIGIN
Alignment Scores:
Pred. No.: 1.79e-62 Length: 791
Score: 912.00 Matches: 176
Percent Similarity: 82.30% Conservative: 10
Best Local Similarity: 77.88% Mismatches: 39
Query Match: 24.39% Indels: 2
Gaps: 0
US-09-821-883-2 (1-690) x B1154872 (1-791)
Qy 322 TyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArgSer 343
Db 111 TTCTTCTCCGACGACCCGCTGAGTCTGAGGACAGACGCCGACGACGCGGACG 170
Qy 342 SerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGlnuGlu 361
Db 171 TCGTCGGCCAGGAGTGGCGGTGAGCTGACACTGGGCTGGAGCCCTCGGAAGAGAG 230
|||||

```

Oy 362 AlAProArGSePrOleuAlAProSeRGlUGlYAlaGlySerAspValPheAspGlyAsp 381
    |||||
Db 231 CCCCCAGATCTCCACGTGGCTCCCTCCGAGGGGCTGGCTCCGATGTTGATGATGAC 290
Oy 382 LeuGlyMetGlyAlAAlaLySGLyLeuGlnSerLeuProThrHisAspProSePrOleu 401
    |||
Db 291 CTGGCACTGGGGGTACCAAGAGACTGCAGAGCTCTCTCCACATGACCTCAGCCCTCTA 350
Oy 402 GlnArGTYrSerGluAspProThrValProLeuProSeRGlUThrAspGlyTYrValAla 421
    |||||
Db 351 CAGGGTAGACAGTAGATCCACATCTACCTGCCCCCGAGACTGATGGCTAGCTTGTCT 410
Oy 422 ProLeuThyGlySerProGlnProGluTYrValAsnGlnProAspValArGProGlnPro 441
    |||||
Db 411 CCCCCTGGCCGACACCCCGCCGAGCTATGTGAACACAGAGGTTGGCCCTCAGTCT 470
Oy 442 ProSeRProArGglUGlyProLeuProAlAAlaArGProAlaGlyAlaThrLeuGluArG 461
    |||||
Db 471 CCTTGAACCCAGAGGGTCTCCGCTCCCATCCGACCTGCTGCTACTGTAGAAGA 530
Oy 462 AlAlyThrLeuSeRProGlyLySAsnGlyValValLySAspValPheAlaPheGlyGly 481
    |||||
Db 531 CCCAAGACTCTCTCTCTGGGAAAAATGGGGTTGTCAAAAGACGT-TTTCCTTGGGGGT 589
Oy 482 AlAValGluAsnProGluTYrLeuThrProGlnGlyAlaAlaProGln-ProHisPr 501
    |||||
Db 590 GCTGTGGAGAACCTGTAACTTAGCACCCAGACAGACGACTGCTCAGACCCCAAC 649
Oy 501 oProProAlaPheSeRProAlaPheAspAsnLeuTYrTYrTPAspGlnsPrProGln 521
    |||||
Db 650 TTCTCTGCTCTTACGCGACGCTTGTGAACCTCTATTACTGGGACCAAGAACTATCGGA 709
Oy 521 uArGlyAlaProProSeRThrPheLySGLYThrProThrAlaGluAsnProGluTYrLe 541
    |||||
Db 710 GCAAGGCTCCACCAAGTACCTTTGAAGGAGCCCAATGACAGAAACCTGTAGTACT 769
Oy 541 uGlyLeuAspValPro 546
    |||||
Db 770 AGGCTGTGATGTGCCA 785

RESULT 10
LOCUS BM790293 491 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0070099 S22SNU16 Homo sapiens cDNA clone S22SNU16-3-A11 5',
            mRNA sequence.
ACCESSION BM790293
VERSION BM790293.1 GI:19138525
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 491)
AUTHORS Kim,N.S., Hahn,Y.J., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,Y.E., Sohn,H.Y., Kim,U.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 3 row: A column: 11
            High quality sequence stop: 491.
FEATURES
    source
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="S22SNU16-3-A11"

```

```

/clone_lib="S22SNU16"
/sex="F"
/tissue_type="Ascites"
/cell_type="Lymphoblast-like"
/cell_line="SNU-16"
/lab_host="DH10B"
/note="Organ: Stomach; Vector: pT73-Pac; Site:1; ECORI.
Site_2: NotI; The S22SNU16 library was contributed by the
Soares laboratory and it was constructed as described by
Bonaldi, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. RNA was prepared from harvested
cells of SNU-16 culture. SNU-16 cell was obtained from
Korean Cell Line Bank (KCLB). SNU-16 was established from
ascitic fluids of Korean patients by Park J.G. et al.
(1990), Cancer Res 50: 2773-2780."
BASE COUNT 100 a 179 c 130 g 82 t
ORIGIN
Alignment Scores:
Pred. No.: 7,31e-62 Length: 491
Score: 901.00 Matches: 163
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24,10% Indels: 0
DB: 14 Gaps: 0

US-09-821-883-2 (1-690) x BM790293 (1-491)
Oy 378 PheAspGlyAspLeuGlyMetGlyAlaAlaLySGLyLeuGlnSerLeuProThrHisAsp 397
    |||||
Db 1 TTGTATGTGATACCTGGGAATGGGGGAGCCCAAGGGGCTGCAAGCTCCACACATGAC 60
Oy 398 ProSeRProLeuGlnArGTYrSerGluAspProThrValProLeuProSeRGlUThrAsp 417
    |||||
Db 61 CCCAGCCCTTACAGCGGTACAGTAGAGACCCACATACCTGCCCCCTGTGAGACTGAT 120
Oy 418 GlyTYrValAlaProLeuThrCysSerProGlnProGluTYrValAsnGlnProAspVal 437
    |||||
Db 121 GGCTACCTTCCCTCCCTGACCTCAGCCCTCAGCTGTAATGTGAACCAAGCAAGAGT 180
Oy 438 ArGProGlnProProSeRProArGglUGlyProLeuProAlAAlaArGProAlaGlyAla 457
    |||||
Db 181 CGGCCCCAGCCCTTGGCCCCGAGAGGCCCTGTGCTGCTGCCCCAGCTGCTGCTGCC 240
Oy 458 ThrLeuGluArGAlaLySGLYThrLeuSeRProGlyLySAsnGlyValValLySAspValPhe 477
    |||||
Db 241 ACCTGTGAAGAGGGCAAGACTCTCTCCCAAGGAGAAATGGGGTGTCAAAAGAGCTTTT 300
Oy 478 AlAlyPheGlyAlaValAlaGluAsnProGluTYrLeuThrProGlnGlyAlaAlaPro 497
    |||||
Db 301 GCTTTGGGGGTCCCGGAGAAACCCGAGTACTTGAACCCACAGGAGAGAGCTGCCCT 360
Oy 498 GlnProHisProProProAlaPheSeRProAlaPheAspAsnLeuTYrTYrTPAspGln 517
    |||||
Db 361 CAGCCCAACCTCTCTCTGCTTCAAGCCAGCTTGACAACTCTATTCTGGAGACAG 420
Oy 518 AspProProGluArGlyAlaProProSeRThrPheLySGLYThrProThrAlaGluAsn 537
    |||||
Db 421 GACCACACAGAGGGGGGCTCACACACACCTTCAAGAGACACCTACGACAGAAAC 480
Oy 538 ProGluTYr 540
    |||||
Db 481 CCAGAGTAC 489

RESULT 11
LOCUS B1649877 998 bp mRNA linear EST 12-SEP-2001
DEFINITION 603296516F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5337435 5',
            mRNA sequence.
ACCESSION B1649877
VERSION B1649877.1 GI:15564113
KEYWORDS EST.
SOURCE house mouse.

```



```
/clone="NT2RM2001211"
/clone_lib="NT2RM2"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/notes="Vector: pME18SF13; mRNA from uninduced NT2 neuronal
precursor cells"
BASE COUNT 137 a 224 c 198 g 123 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 4,31e-60 Length: 685
Score: 881.00 Matches: 170
Percent Similarity: 91.89% Conservative: 0
Best Local Similarity: 91.89% Mismatches: 8
Query Match: 23,56% Indels: 7
DB: Gaps: 1

US-09-821-883-2 (1-690) x AW123871 (1-685)
QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHis 337
DB 137 TGTCCA-----GACCCCTGCCCCGGGGCGCTGGGGCATGTCCACCAC 178
QY 338 ArgHisArgSerSerThrArgSerGlyGlyValAspLeuThrLeuGlyLeuGluPro 357
DB 179 AGGCACCGCAGCTCATCTACAGAGAGTGCGGGGAGCCTGACACTAGGGCTGGAGCCC 238
QY 358 SerGluGluGluAlaProArgSerProLeuAlaProSerGlyAlaGlySerAspVal 377
DB 239 TCTGMAAGAGGAGGGCCCGACAGTCTCCACCTGCAACCCCTCGAAGGGGCTGGCTCGATGA 298
QY 378 PheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuProThrHisAsp 397
DB 299 TTTCATGATGATCACTGGGAAGGGGGGAGCCCAAGGGGCTGCAAAAGCCCTCCACACATGAC 358
QY 398 ProSerProLeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAsp 417
DB 359 CCCAGCCCTTACAGCCGTCATGAGGAGACCCACACTACCTGCTGAGACTGAT 418
QY 418 GlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAlaGlnProAspVal 437
DB 419 GGCTAGCTGGCCCCCTGACCTGAGCCGCCAGCTGAATATGTGAACAGCCAGATGTT 478
QY 438 ArgProGlnProProSerProArgGlyGlyProLeuProAlaAlaArgProAlaGlyAla 457
DB 479 CGGCCCCAGCCCTTGGCCCCGAGAGGGCTCTGCTCTGCTGCCACCTGCTGGTGC 538
QY 458 ThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPhe 477
DB 539 ACTCTGGAAGAGGCGCAAGACTCTCTCCCGAGGAGAAATGGGGTCTGCAAAAGACGTTTTT 598
QY 478 AlaPheGlyGlyAlaValGluAsnProGluTyr-LeuThrProGlnGlyGlyAlaAlaPro 497
DB 599 GCGTTTGGGGTGGCCGCGAGAACCCCGAGTACTTGTACACCCCGANGAGAGAACTGCCCC 658
QY 497 OGlnProHisPro 501
DB 659 CTTAAAGCCACCC 671

RESULT 13
AW701942 609 bp mRNA linear EST 18-APR-2000
LOCUS uq93802.y1 NCI_CGAP_Mam10 Mus musculus cDNA clone IMAGE:2938635 5'
DEFINITION similar to gb:M1130 ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE
PRECURSOR (HUMAN); mRNA sequence.
ACCESSION AW701942
VERSION AW701942.1 GI:7586086
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 609)
```

```
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Jeffrey Green M.D.,
Gilbert Smith, Ph.D., William Muller, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
image.lnl.gov/image/html/iresources.shtml
MGI:1051055
Seq primer: -40RP from G1bco
High quality sequence stop: 512.
FEATURES
Location/Qualifiers
1..609
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:2938635"
/clone_lib="NCI_CGAP_Mam10"
/tissue_type="pooled mammary gland tumors"
/lab_host="Dh10B (phage-resistant)"
/notes="Organ: mammary gland; Vector: pT73D-Pac (Pharmacia
) with a modified polylinker; Site.1: NotI; Site.2: EcoRI;
1st strand cDNA was prepared from mRNA obtained from
pooled mammary gland tumors with a Not I - oligo(dT)
primer [5'
TGTACCAATCTGAAGTGGAGGCGGACATGATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldi."
BASE COUNT 125 a 195 c 159 g 127 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 4e-59 Length: 609
Score: 866.00 Matches: 163
Percent Similarity: 85.07% Conservative: 8
Best Local Similarity: 81.09% Mismatches: 30
Query Match: 23,21% Indels: 0
DB: Gaps: 0

US-09-821-883-2 (1-690) x AW701942 (1-609)
QY 346 SerGlyGlyAspLeuThrLeuGlyLeuGlnProSerGluGluAlaProArgSer 365
DB 1 AGTGGAGGTGTGAGCTGACACTGGGCTGAGCCCTCGGAAGAGGCCCCCAGATCT 60
QY 366 ProLeuAlaProSerGlyGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGly 385
DB 61 CCACGTGCTCTCCGGAAGGGGCTGGCTCGAGTGTATTATGTGACCTGCGCAATGGGG 120
QY 386 AlaAlaGlyGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSer 405
DB 121 GTAAACCAAGGGCTGCAAGACCTCTCTCCACATGACCTCAGCCCTTACAGCGGTACAGC 180
QY 406 GluAspProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCys 425
DB 181 GAGGAGCCCGCATTTACTCTGCCCCCGAGACAGATGATGCTGCTGCCCTGGGCTGC 240
QY 426 SerProGlnProGluTyrValAlaGlnProAspValArgProGlnProProSerProArg 445
DB 241 AGCCCCAGCCCGATATGTGAACCAATCAAGAGTTCAGGCTCAGGCTCTTAACCCCA 300
QY 446 GluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeu 465
```


Db	301	GAGGGTCCTCTGCGCCGTCGTCGGCCCTGCTGCTACTCTAGAAAGAACCCAGACCTCTC	360
QY	466	SerProGlyLyAsnGlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsn	485
Db	361	TCTCTCTGGAGAAAGGGGCTGTCTAAACACGTTTTTGTCCCTCCGGGGCTGTGTGAGAAC	420
QY	486	ProGluTyrLeuThrProGlnGlyValAlaIleProGlnInProHisProProProIleAlaPro	505
Db	421	CCTGATACTTATGATACCGAGAGAGAGGACCTGCTCTCCGCCCCACCTCTTCTCTGCTTC	480
QY	506	SerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProProProGluArgGlyAlaPro	525
Db	481	AGCCACGCTTTGACACACCTCTATATACGNGACGAGAACTCATCGAGACAGAGNGCTCCA	540
QY	526	ProSerThrPheIleGlyGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspVal	545
Db	541	CCAACTACTCTTGAAGGAGACCCCTCCTCAGAGAACCTGTAGTACTTGGCTTGATGTA	600
QY	546	Pro	546
Db	601	CCT	603
RESULT 14			
LOCUS	BG283493	621 bp	mRNA
DEFINITION	602407782p1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4519652 5',		
ACCESSION	BG283493		
VERSION	BG283493.1	GI:13033480	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 621)		
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DRP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLM10416 Row: d Column: 21 High quality sequence stop.		
FEATURES			
source	1..621	Location/Qualifiers	
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4519652"		
	/clone_1lb="NIH_MGC_91"		
	/tissue="type="adenocarcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: prostate; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."		
BASE COUNT	131 a 207 c 178 g 104 t	1 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	7.78e-59	Length:	621
Score:	864.50	Matches:	174
Percent Similarity:	91.67%	Conservative:	2
Best Local Similarity:	90.62%	Mismatches:	12
DB:	12	Indels:	1
	12	Gaps:	7

US-09-821-883-2 (1-690) x BG283493 (1-621)

OY 371 GUGUyAlAGlySerAsPvAlAPheASpGLysPlenGluMetGlYalAALySgilyLeu 390
|||||
Db 2 GAAGGGCGTGCCTCCGAATGTATTGTATGTACCTGGCAAAAGGGGCCAACGAGGGCCTG 61

OY 391 GinserleUpProThrHisASPProSeSProLengInAgTyrsERGuASPProthrVal 410
|||||
Db 62 CAAMACCtCCCCACAATGACCcCCCAGccctTCAGCcGGrCACAGTGAGAGAcCCCACTA 121

OY 411 PROLEUPROSErgLUThrASpGLYTyrVALAAProleuthRcySSerPrOglINPROGU 430
|||
Db 122 CCCCGGCCCTCGACTCATGATGGTAGCTGCCcCCCGCTGACCTGCAGCCcCCAGCCTGA 181

OY 431 TyrrVAlAnsglnPRoASpVAIArgPRoglinPRoPSerPRoArGguLyprOLEUPRo 450
|||
Db 182 TATGTGAACCAAGCCAGATGTTGGCCCGCACGCCcCTTCGGccccGAGAaggcccTTCGCT 241

OY 451 ALaAtARCPROLAGlIALAthrLeugLuIrFgaLatYsrThrLeusErPGLYlysAn 470
|||
Db 242 GTGCGCCGACCTGCTGGrGCCACTGTGGAAAAGGCCCAAGACTCTCCCCAGGAGAAT 301

OY 471 GlyvalValysaspvaIpheaIApheGIglylaVaIlguaSnPRoglUTyrLeUTHr 490
|||
Db 302 GGgggcGTCAAAGAGGTtttgCCtTYTGggGGGTgccGTGGAGMAcccgagActTYTGACA 361

OY 491 ProGlnglyGYAlALAIProgLn-ProHISprothroPraIApheserPROLIAPhas 510
|||
Db 362 CCCC-A-GGAGGAGCTNGCCCTCACGCCcCAACCCCTTCCTCGCTTCACCCACGCTTGA 420

OY 510 PASNLeutyRTYTRPASgLIAsPPrcProPgIUarGgLYAlAProprosErtThrPhely 530
|||
Db 421 CAacctCrITnTrCTGGACACAG -CCAACAAGGGGGGCGCTCCAccCAAGAC-TTCAA 478

OY 530 sGLYThrProTHraIgUAsnPRogLUtyrLeungILyLeuSpvaIProAlA----- 547
|||
Db 479 AGGGACACCTACGGAGAGAACCCAGAGTAactGTGGGTGGAGCaGTGaAGCAG 538

OY 548 -AlAlaIProAIARGseRProseRProserTHR 558
|||
Db 539 AAGCAAGTCGCGAGAGCCCTGATGTGTCTCA 572

RESULT 15
B0711097 887 bp mRNA linear EST 16-JUL-2002

LOCUS B0711097
DEFINITION AGENCOURT_8100479 LupsK1_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6190537 5', mRNA sequence.

ACCESSION B0711097
VERSION B0711097.1 GI:21855994
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapIans
*Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.*

REFERENCE 1 (bases 1 to 887)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: ILAMI3589 row: 1 column: 02
High quality sequence stop: 556.
Location/Qualifiers .
.I..887
./organism="Homo sapiens"

FEATURES
Source

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 7, 2003, 22:19:33 ; Search time 1592 seconds

(Without alignments)
976.055 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLAARASLSLGLF.....EPVCEGAPPPAAAHNNHHH 690

Scoring table:

BLOSUM62	Xgapop 10.0 , Ygapext 0.5
	Xgapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=US-09-821-883-2.p2n.rng -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE_PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -USER=US09821883 @CGN.1.1.396 @runat.28042003.104748.14558 -NCPU=6 -ICPU=3 -NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120 -WAIT_TIMEOUT=30 -THRAVS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	N.Geneseq.101002:*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001.DAT:*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3739	100.0	2070	22 AAD21565	Human HER500-hgm-c
2	3473.5	92.9	2091	22 AAD21567	Human HER500-rgm-c
3	2974	79.5	1665	22 AAD21564	Human HER500 fusio
4	2959.5	79.2	1692	22 AAD21566	Human HER500 fusio
5	2423	64.8	4530	16 AAT01585	Her-2/neu (ERBB2)/c
6	2423	64.8	4530	18 AAT71253	Human HER2 gene.
7	2423	64.8	4530	21 AA260815	Nucleotide sequenc
8	2423	64.8	4530	22 AAD19731	Human tyrosine kin
9	2423	64.8	4530	22 AAD19731	Human HER2-neu SEQ
10	2423	64.8	4530	24 ABR83918	Human CDNA differe
11	2422	64.8	9274	24 ABR14057	Human HER2 (ETB2)
12	2418	64.7	4472	21 AA141812	CDNA encoding the
13	2418	64.7	4473	20 AA231071	HER-2 nucleic acid
14	2418	64.7	4473	24 AAD38904	Human Her-2 DNA.
15	2415	64.6	9274	22 AAF24297	HER2 transgene pla
16	2410.5	64.5	3768	21 AA09455	Human heregulin 2
17	2410.5	64.5	3768	21 ABL91709	Human polynucleoti
18	2410.5	64.5	3768	24 ABR14058	Human HER2 (ETB2)
19	2405.5	64.3	3768	17 AAT40739	Human HER-2/neu oncogene
20	2405.5	64.3	3768	20 AA01912	Human HER-2/neu on
21	2405.5	64.3	3768	22 AAD23392	Human HER-2/neu pr
22	2405.5	64.3	3768	24 AAD22743	Human Her-2/neu pr
23	2405.5	64.3	3768	24 ABA92250	Human Her-2/neu CD
24	2405.5	64.3	3768	24 ABR10730	Human Her-2/neu DN
25	2336.5	64.1	4299	14 AA046083	Sequence encoding
26	2334.5	62.4	3678	14 ABR66207	HER2-GM-CSF immuno
27	2258.5	60.4	2385	18 AAT72725	CDNA encoding huma
28	2188.5	58.5	1437	22 AAD21568	Human HER300-rgm-c
29	2130	57.0	2781	24 ABA92253	Human Her-2/neu ex
30	2129	56.9	2763	24 ABA92252	Mouse Her-2/neu ex
31	2092.5	56.0	3600	21 AA089736	Human HER-2/neu co
32	1968	52.6	3955	16 AAT01590	Rat Her-2/neu promotor.
33	1968	52.6	3955	21 AA089753	Rat Her-2/neu prot
34	1961	52.4	3771	21 AA089753	Mouse Her-2/neu cd
35	1961	52.4	3771	22 ABA92210	Nucleotide sequenc
36	1961	52.4	3771	24 ABA92251	Mouse Her-2/neu cd
37	1680.5	44.9	2871	21 AA250586	DC83Fv-erbB2EC fu
38	1614.5	43.2	1872	11 AA006828	Extracellular port
39	1587	42.4	867	22 AAD21571	Mature human HER-2
40	1199.5	32.1	1191	22 AAD21573	Human NY-ESO-1C-HE
41	1188	31.8	1806	24 AAD32745	Human CDNA for the
42	1183	31.6	1755	24 AAD32746	Human CDNA for the
43	1183	31.6	1767	24 AAD32744	Human CDNA for the
44	1183	31.6	1773	24 AAD32747	Human CDNA for the
45	1182	31.6	651	22 AAD21572	Mature human HER-2

ALIGNMENTS

RESULT 1
AAD21565
ID AAD21565 standard; DNA; 2070 BP.
XX AAD21565;
DT 28-JAN-2002 (first entry)
XX
DE Human HER500-hgm-CSF fusion DNA construct.
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW HER500-hgm-CSF fusion DNA; ds.
XX

21 LeuLeuPhePheTrpLeuAspArgSerValIleuAlaIysGluLeuAlaArgGlyAla 40
61 CTGCTTTTTCCTGGCTGACCGAAGTGTACTAGCAAGAGTGGCGGCGGCGCGG 120
41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
121 TCGACCAAGGTGTGACCGGCGACAGACATGAAGTCGGGCTCCCTGCCAGTCCGAGACC 180
61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnIysAsnLeuGlu 80
181 CACCTGGACATGCTCCGCCACCTCTACACAGGGCTCCAGAGTGTGTGACGGAACCTGGAA 240
81 LeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIValGln 100
241 CTCACCTTACCTGCGCCACCAATGCGAGCTGTCTTCGACAGAAATCCAGAGGTGGAG 300
101 GlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
301 GGCTACGTGCTCATGTGCTCACACCAAGTAGAGGCTCCCTCCAGAGGCTGGGATTT 360
121 ValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnIysAsp 140
361 GTGCGAGGACACCCAGCTCTTTGAGGACAATAGCTGCGCGCTGCTGACAAATGGAGAC 420
141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgIleuGln 160
421 CCGGTGAACAATACCAACCCCTGTCCAGGGGCGCTCCAGAGGCGCTCGGGAGCTGGAG 480
161 LeuArgSerLeuThrGluIleLeuIysGlyValLeuIleGlnArgAspProGlnLeu 180
481 CTTCGAACCTTCACAGACATCTGAAAGAGGGGTCTTGATCCAGCGGACCCCAAGCTC 540
181 CysTrpTrpGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
541 TGTATCCAGGACACATTTTGTGGAGAGCATCTTCCACAAGAACACAGCTGCTCTC 600
201 ThrIleuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
601 ACACGTGATGACACCAACCGCTCTGCGGCTGTCCACCCCTGTCTTCGATGTGAAGGCG 660
221 SerArgCysTrpGlyGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
661 TCCCGCTCTGTGGGAGAGATTGTGAGATTGTGAGACCTGAGCGGACAGTGTCTGTCC 720
241 GlyGlyCysAlaIleArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
721 GTGCGCTGTGCCCTGCCAGGGGCGCTGCCACTGACTGCTGCCATGAGAGAGTGTCT 780
261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
781 GCGGCTCAGGGGCGCCCAACAGCAGCTCTACTGCTGCGCTGCTCCACTTCAACACAGT 840
841 GlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGluSer 300
901 MetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp 320
301 MetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp 320
901 ATGCGCAATCCCGAGGGCGGTATACATTCGGCGCAGCTGTGTGACGTCTGTCCAC 960
321 AsnTrpLeuSerThrAspValAlaIysSer-----GlyAla 331
961 AACATACCTTTCTACGAGAGTGGATCCGCTAGCATCATTAATTTTCGAGAACTGGGCGCT 1020
332 GlyLeuMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyIysPheLeu 351
1021 GGGGGCATGTGTCCACACAGGCGACGCTCATCTACACAGAGTGGGGGGAGCTTG 1080
352 ThrLeuGlyLeuGluProSerGluGluGluAlaProAlaArgSerProLeuAlaProSerGlu 371
1081 ACACGTAGGGCTGTGAGGCTCTGTGAAGAGAGGCCGCCACAGTCTCCACCTCCCAA 1140
372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaIysGlyLeuGln 391

1141 GGGGCTGGCTCCGATGTATTGTATGTGTGACCTGGGAATGGGGGCGCAAGGGGCTGGCA 1200
392 SerLeuProThrHisAspProSerProLeuGlnArgTrpSerGluAspProThrValPro 411
1201 ACCCTCCCAACATGACGCCCGCCCTCTACAGCGGTACAGTGAAGAGACCCCACTATACC 1260
412 LeuProSerGluThrAspGlyTrpValAlaProLeuThrCysSerProGlnProGluTrp 431
1261 CTGCGCTGTGAGACTGATGGCTACTGCTCCCTGACTGTGACGCGCCACGCTGAAATAT 1320
432 ValAsnGlnProAspValArgProGlnProProSerProAlaArgGluGlyProLeuProAla 451
1321 GTGAACAGGACGATGTGGGCGCCAGCCCTTGTGCGCCGAGAGGGCGCTGTGCTGCT 1380
452 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 471
1381 GCCGACCTGCTGTGGTCCACTCTGGAAAGGCGCAAGACTCTCTCCCAAGGAAATGGG 1440
472 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTrpLeuThrPro 491
1441 GTCTCAAAAGAGCTTTTGGCTTTGGGGGTGGCGGAGAACCCCGAGTACTGACACCC 1500
492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511
1501 CAGGAGAGAGCTGCCCTCAGGCCACCTCTCTGCTTCAGGCCAGCTTGCAACAC 1560
512 LeuTrpTrpTrpAspGluAspProProGluArgGlyAlaProProSerThrPheLysGly 531
1561 CTCTATTACTGGGACACAGAGCCACACAGCGGGGCTCCACCCAGACCTTCAAAAGG 1620
532 ThrProThrAlaGluAsnProGluTrpLeuGlyLeuAspValProAlaAlaProAla 551
1621 ACACCTACGGAGAGAACCCAGAGTACTGGGTGTGAGAGTGTGAGCGGCGGCCCAACC 1680
552 ArgSerProSerProSerThrGlnProThrGlnHisValAsnAlaIleGlnGluAlaArg 571
1681 CGCTACCCCAACCCCTGTACCGCGGCTGTGAAGCATGTGATGATGATCAAGAAAGCTGTG 1740
572 ArgLeuLeuAsnLeuSerArgAspThrAlaAlaGluMetAsnGlnIleValGluValIle 591
1741 AGCCTCCAAATGACATCGTGTGTGAGAAACGAAGAAAGAACAGAGTACATCATC 1800
592 SerGluMetPheAspLeuGlnGluProThrCysLeuGlnIleThrArgLeuGluLeuTrpLys 611
1801 TCTATGAGTTCTGCTCATCAGAGCGGCACATGTGTGTGAGACCCCGCTGAACTGTATCAAG 1860
612 GlnGlyLeuArgGlySerLeuThrLysLeuIysGlyProLeuThrMetMetAlaSerHis 631
1861 CAGGCTTACGGGGCGCACTCAACCACTCAATGCGCTTGTACATATAGCCAGCCAC 1920
632 TyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGlnIleIleThrPhe 651
1921 TACAGAGCAACTGCGCTCCAAACCCCGAAACTGACTGAAATATGAAGTCAACACCTTT 1980
652 GluSerPheLysGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGlu 671
1981 GAGATTTTCATAAAGAACCTTAAAGGCTTGTGTGATATCCCTTTGATGCTGTGGAG 2040
672 ProValGlnGluGlyAlaProProProProAlaAlaHisHisHisHisHis 690
2041 CCGGTCCGAAGAGGCGGCCACCCCGCGG-----GGCATCAACATCAACATCAC 2091

RESULT 3
AAD21564
ID AAD21564 standard; DNA; 1665 BP.
AC AAD21564;
XX 28-JAN-2002 (first entry)
XX Human HER500 fusion DNA construct.
XX

Oy	401	LeugAlaargTyrSerGluAspProThrValProleuproSergLutHrAspLysIrrVal	420
Db	1201	CTACAGCGGTACTGGAGGCCACCACAGTACCCCTCCTTGAGACTGATGGTACGTT	1260
Oy	421	AlaProLeuthrCySserProGlnProGluTyrValAsnGlnProAspValArpProGln	440
Db	1281	GCCCCCTGACCTGCACCCCACGCCGGAATATGTGAACACCAAGATGTTGGCCCCAG	1320
Oy	441	ProProSerProAlaArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu	460
Db	1321	CCCCCTTCGCCCGAAGAGGGCCCTCTGCCTGCCCGCCGACCTGTGGTCCACTCTGGAA	1380
Oy	461	ArgAlaIalYsrThleuSerProGlyLySasnGlyValIvalLySaspValPheAlaPheGly	480
Db	1381	AGGGCCAGAACTCTCTCCCGAGGAAGATGGGTCGTCAAAGACGTTTTTGCTTTGGG	1440
Oy	481	GlyAlaValGluAsnProGluTyrLeuthrProGlnGlyGlyAlaAlaProGlnProHis	500
Db	1441	GGTCCCGTGGAGAAGCCCCGAGTACTTGACACCCACGAGGAGAGCTGCCCTCAGCCAC	1500
Oy	501	ProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro	520
Db	1501	CCTCTCTCTGCTTCACCCACGCTTCGACAACTCTATTACTGGAGCACGAGACCCACCA	1560
Oy	521	GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr	540
Db	1561	GAGCGGGGGGCTCCACCCACACCTTCAAAAGGACACCTACGCGACAGAACCCAGAGTAC	1620
Oy	541	LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSer	557
Db	1621	CTGGGTCTGGACGTGCCGACGCGGCC-----GCACA-TCACCATCACATCA	1664
RESULT 4			
ID	AAD21566	standard; DNA; 1692 BP.	
XX	AAD21566;		
DT	28-JAN-2002	(first entry)	
DE	Human HER500 fusion DNA construct comprising OVA-derived octapeptide.		
XX			
KW	Immunostimulatory fusion protein; IFP; antigen component; therapy;		
KM	immunostimulatory component; T-cell mediated immune response; DC;		
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;		
KM	PAP proteol; Ala Arg linker; membrane distal extracellular domain;		
KW	membrane distal intracellular domain; C-terminal tag; human; OVA;		
KM	HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion DNA; ds.		
XX			
OS	Chimeric - Homo sapiens.		
OS	Chimeric - Unidentified.		
XX			
FH	Location/Qualifiers		
FT	1..1692		
FT	/tag= a		
FT	/product= "Human HER500 fusion protein construct		
FT	comprising human PAP signal sequence, mature PAP		
FT	protein, an Ala Arg linker, human HER-2 signal		
FT	sequence, mature HER-2 membrane distal extracellular		
FT	domain, an Ala linker, an ovalbumin (OVA)-derived		
FT	immunodominant octapeptide, HER-2 membrane distal		
FT	intracellular domain and a C-terminal tag"		
FT	/note= "CDS does not include stop codon"		
FT	/partial		
XX			
PN	WO200174855-A2.		
XX			
PD	11-OCT-2001.		
XX			
PE	30-MAR-2001: 2001WO-US10515.		
XX			
PR	30-MAR-2000: 2000US-193504P.		

XX	(DEND-) DENDREON CORP.		
XX			
XX	Laus R, Vidovic D, Gradidis T;		
XX	WPI: 2001-662965/76.		
DR	P-PSDB: AAEL3110.		
XX			
PT	An immunostimulatory fusion protein comprising the intracellular domain		
PT	of HER-2 and an antigen elicits an immune response to the antigen and		
PT	is useful for the treatment of associated cancer associated		
XX			
PS	Example 3; Page 28; 59pp; English.		
XX			
CC	The invention relates to immunostimulatory fusion proteins (IFP) and		
CC	nucleic acid molecules encoding such proteins. The IFPs comprise a		
CC	polypeptide antigen component and an immunostimulatory component derive		
CC	from the intracellular domain of HER-2 protein which is effective to		
CC	elicit a protective dendritic cell (DC)-induced T-cell mediated cellular		
CC	immune response to the antigen. IFP or superactivated dendritic cells		
CC	are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer		
CC	associated with a particularly antigen. The present sequence is HER500		
CC	fusion DNA construct which comprises DNA molecules encoding human PAP		
CC	signal sequence, mature PAP protein, an Ala Arg linker, human HER-2		
CC	signal sequence, mature HER-2 membrane distal extracellular domain,		
CC	an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,		
CC	HER-2 membrane distal intracellular domain and a C-terminal tag.		
XX			
SO	Sequence 1692 BP; 338 A; 573 C; 470 G; 311 T; 0 other;		
	Alignment Scores:		
	Pred. No.: 9.3e-122	Length: 1692	
	Score: 2959.50	Matches: 554	
	Percent Similarity: 97.88%	Conservative: 0	
	Best Local Similarity: 97.88%	Mismatches: 1	
	Query Match: 79.15%	Indels: 12	
		Gaps: 2	
US-09-821-883-2 (1-690) x AAD21566 (1-1692)			
QY	1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe		20
Db	1 ATGAGAGCTGCACCCCTCTCTCTGCGCAGGGGAGCAACCTTAAGCTTGGCTTCTTT		60
QY	21 LeuLeuPhePhePheTrpLeuAspArgSerValLeuAlaValGluLeuAlaArgValAla		40
Db	61 CTGCTTTTCTTCTGGCTGAGCCGGAAGTGTACTAGCAAGAGATGGTGGCGGGGCGCG		120
QY	41 SerThrGlnValCysThrGlyThrAspMetCysLeuAlaGlyLeuProAlaSerProGluThr		60
Db	121 TCGAACCAAGTGTGACGGCGCAGACAGACTGAAGCTGGGGCTCCCTGCGACGTCGAGACC		180
QY	61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnGlyAsnLeuGlu		80
Db	181 CACCGGGAACATGCTCCGACCTCTACCAAGGGCTGCCAGGTGGTCACGGAAACCTGGAA		240
QY	81 LeuThrTyrLeuProThrAspAlaSerLeuSerPheLeuGlnAspIleGlnIleValGln		100
Db	241 CTCACCTACTCTGCCACCAATGCCAGCTGTCTTCTCTGACGATATTCAGAGAGGTGCAG		300
QY	101 GlyTyrValLeuIleAlaHisAspGlnValAlaArgGlnValProLeuGlnArgLeuArgIle		120
Db	301 GGCTACGTCGTCATGCGCTCACAACCAAGTGAAGGCGAGTCCACATGCAAGAGCTGGGATT		360
QY	121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp		140
Db	361 GTGGGAGGACCCAGCCTCTTGTAGGAGACAATATGCTCCCTGGCTGTAGACAATGTGAAC		420
QY	141 ProLeuAsnAsnThrThrProValThrCylAlaSerProGlyGlyLeuArgIleGlnGln		160
Db	421 CCGGTGAACATATACCAACCCCTCTGTACAGGGGCTCCACGAGAGCGCTGGGAGACTGCAG		480
QY	161 LeuArgSerLeuThrGluIleLeuGlyGlyValValLeuIleGlnArgAsnProGlnLeu		180


```

Db      481  CTTCAACCCCTCACAGATCTTTGAAGAGGGGCTTGATCCAGCCGAGACCCACAGCTC
Oy      181  CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAspGluLeuAlaLeu
Db      541  TGTATCCAGGACACATTTTGGAGAGCATCTTCCAAAGAACACAGACTGGCTCTGC
Oy      201  ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly
Db      601  AACATGATAGAACACACCGCTCTCGGGCTGCCACCCCTGTTCCGATGTGTAAAGGC
Oy      221  SerArgCysTTPGlyGlnSerSerGluAspCysGlnSerLeuThrArgThrValCysAla
Db      661  TCCCGCTCTGGGGAGAGATTCTAGAGATTCTCAGACCTGACCCGACCTCTGTCTGC
Oy      241  GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla
Db      721  GGTGGCTGTGCCCTGCCAAGGGGCCACTGCCCTGACTGCTGCTCATGAGCAGTGTGCT
Oy      261  AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer
Db      781  GCCGGCTGCAGGGGCCACAGACATCTGACTGCTGGCTGCTGCCTCAGCTTACACAGT
Oy      281  GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGlnSer
Db      841  GGCATCTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Oy      301  MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr
Db      901  ATGCCCAATCCGAGGGCGGTATACATTCGGCGCAGCTGTGCTGCTGCTGCTGCTGCT
Oy      321  AsnTyrLeuSerThrAspValGlySer-----GlyAla 331
Db      961  AACTACCTTTCTAGAGAGTGGGATCCCTACATCATTAATTTCGAGAACTTGGCGCT
Oy      332  GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyLysAlaLeu
Db      1021  GGGGCGATGTGCTCCACAGCGCCAGCGCATCTATCTACAGAGTGGCGGTGGGACCTG
Oy      352  ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu
Db      1081  ACACTAGGGCTGGAGCCCTCTGTAAAGAGGCCCCAGGTCTCCACTGGCACCTCCGAA
Oy      372  GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln
Db      1141  GGGCGTGGCTCGATGTATTTGATGTGATCGTGGGAATGGGGGAGCCAAAGGGCTGCA
Oy      392  SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro
Db      1201  AGCTTCCCCACACATGACCCCTCTACAGCGGTACAGTACAGAGAGCCCCACACTACCC
Oy      412  LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlyTyr
Db      1261  CTGGCCCTGTGACATGATGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
Oy      432  ValAsnGlnProAspValAlaArgProGlnProProSerProArgGlyGlyProLeuProAla
Db      1321  GTGAACACAGCCAGATGTGGGCCAGGCCCTTGGCCCGAGAGGGCCCTTGGCTGCT
Oy      452  AlaArgProAlaGlyAlaThrLeuGluAlaLysThrLeuSerProGlyLysAsnGly
Db      1381  GCCCGACTGTGTGGCCACTCTGAAAGGGCCAAAGACTCTCTCCCAAGAGAAATGGG
Oy      472  ValValLysAspValPheAlaPheGlyGlyAlaValGluAspProGluTyrLeuThrPro
Db      1441  GTCTCAAGAGAGCTTTTGGCTTGGGGGTGGCGTGGAGAACCCGAGTACTTGGACACC
Oy      492  GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn
Db      1501  CAGGGAGAGAGCTGCCCTCAGCCACCTCTCTCTGCTTACGCCAGCTTTCAGCAAC
Oy      512  LeuTyrTyrTTPAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly

```

```

Db      1561  CTCTATTACTGGGACAGACCACAGAGCGGGGGCTCCACCCAGCAGCTTCAAGGG 1620
Oy      532  ThrProThrAlaGluAspProGluTyrLeuGlyLeuAspValProAlaAlaProAla 1551
Db      1621  ACACCTACGGCGCAGAGAACCCAGAGTACTGGGTCTGGACGTGCCAGCGGCC-----GCA 1674
Oy      552  ArgSerProSerProSer 557
Db      1675  CA-TCCACATCAGCATCA 1691

```

RESULT 5

AAAT01585 standard; DNA; 4530 BP.

AAAT01585;

20-APR-1996 (first entry)

Her-2/neu (ERBB2/c-erbB-2) gene sequence.

Her-2/neu; Erb-B2; c-erbB-2; oncogene; DNA binding protein; HPRF;

Erb-B2 promoter binding protein; tumour enhancer factor;

breast cancer diagnosis; prognosis; antisense oligonucleotide;

retro virus vector; gene therapy vector; ss.

Homo sapiens.

W09528485-A1.

26-OCT-1995.

19-APR-1995; 95WO-US04953.

19-APR-1994; 94US-0229515.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Raziuddin F, Sarkar FH.

WPI: 1995-373800/48.

New purified protein binding to the ERBB2 gene promoter - to induce

cell proliferation, diagnostic of breast cancer, also related

antibodies, nucleic acid, assays and methods for screening

inhibitors.

Disclosure: Page 52-54; 69pp; English.

The Erb-B2 gene is one of the primary genes responsible for the translation of normal breast epithelial cells towards carcinoma in situ and the subsequent development of invasive and metastatic cancer. HPRF (see AAR77093-94), the Erb-B2 promoter binding protein, induces cell division on binding to the promoter. In a method for greater success in early identification and treatment of breast cancer, the initiation step for Erb-B2 gene activity is identified. This method involves determining the presence of HPRF in a biopsy from the subject, where the presence of HPRF (relative to its absence in a normal control) indicates the presence of cancer and a decreased chance of long-term survival. Binding of HPRF to the CC promoter can be inhibited using antisense oligonucleotides or a non-genomic nucleic acid that binds to HPRF; these oligos can be expressed from retro virus or other gene therapy vectors.

Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other;

Alignment Scores:

Pred. No.: 8,15e-98 Length: 4530

Score: 2423.00 Matches: 524

Best Local Similarity: 41.14% Conservative: 5

Query Match: 64.80% Mismatches: 21

DB: 16 Gaps: 4

US-09-821-883-2 (1-690) x AAT01585 (1-4530)

QY 2 ArgAlaAlaProLeuLeuLeuAlaAlaAla-----AlaSerLeuSerLeuGly 17
Db 94 CGCCGCCCTCCAGCCGGGCTCAGCCGGAGCCATGGGGCCGAGCGGATGACACCATG 153
QY 18 PheLeuPheLeuLeuPhePhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
Db 154 GAGCTGGGGGCTTGTGGCCGTG-----GGGCTCCCTCTCCGCTTCCCTCCCTCC 204
QY 38 GluAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 205 GGAGCCCGGAGACACCAAGTGTGCACCGGACAGACATGAAGCTGGCGCTCCCTGCACT 264
QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnGly 77
Db 265 CCCGAGACCCACCTGAGCATGCTCCGGCACTTACAGGGCTCCAGGTGTGCAGAGGA 324
QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGAACTCACCCTACCTGCCCCACCAATGCCAGCTTCTTCAGAGATATCCAG 384
QY 98 GluValGlnGlyTyrValIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
Db 385 GAGGTGACGGGCTACGTGCTCATGTGCTCACAACCAAGTAGAGGCGCTCCCTCAGAGG 444
QY 118 LeuArgIleValaArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
Db 445 CTGGGGATTGTGCGAGGACCCACCTCTTGGAGGACAATGCCCTGGCGCTGCTAGAC 504
QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505 AATGGAGACCGGCTGAACAATACCAACCCCTGTCACAGGGGCTCCCAAGAGGCTCCGG 564
QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValIleGlnArgAsn 177
Db 565 GAGCTGACGGCTGCAAGCTCAGAGATCTGGAAGAGGGGCTTGTATCCAGCGAGAC 624
QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln 197
Db 625 CCCAGCTCTGCTACAGGACACATTTTGGAGAGACATCTTCCACAAGAACACACAG 684
QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
Db 685 CTGGCTTTCACACTGATGACACCAACCGCTCTCGGCTGCGCAACCCCTGTTCCAGTG 744
QY 218 CysLysGlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237
Db 745 TGTAAAGGCTCCCGCTGCTGGGAGAGAGTTCAGAGATTGTGACAGCTGACGCACT 804
QY 238 ValCysAlaGlyGlyCysAlaIaArgCysLysGlyProLeuProThrAspCysHisGlu 257
Db 805 GTTGTGCTCCGCTGCTGCGCTGCGCAAGGGCCACTGCGCACTGCTGCTGCTGCTGCT 864
QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
QY 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
Db 925 AACCAAGTGGCATCTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
QY 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
Db 985 TTTGAGTCCATGCGCAATCCCGAGGGCGGTTATACATTCGGCGCCACAGCTGTGACTGCC 1044
QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
Db 1045 TGTCCCTACAACTACTTCTTACGGAGCTGGATCTGACACCTCTGCTCCCTGCTGCTGCT 1104
QY 329 ----- 329
Db 1105 AACCAAGAGGTGACAGCAGAGATGGAACACAGCGGTGTGAGAACTGACACAGCCCTGT 1164

QY 329 ----- 329
Db 1165 GCCGAGTGTGCTATGCTGTGGCATGAGACATCTGGAGAGAGTGGAGGCGATTACCACT 1224
QY 329 ----- 329
Db 1225 GCCAATATCCAGAGATTGTGCTGCTGCAAGAAGATCTTTGGAGCTTGGCATTTCTGCCG 1284
QY 329 ----- 329
Db 1285 GAGAGCTTTGATGGGGAGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGACCTCAA 1344
QY 329 ----- 329
Db 1345 GTGTTGAGACTGTGAAGAGATCAGAGATTACCTATATCATCTGAGCATGSGGGAGACG 1404
QY 329 ----- 329
Db 1405 CTGCTGACCTCAGCGTCTTCCAGAACCTGCAAGTAATCCGGGAGCGAATTCTGCACAA 1464
QY 329 ----- 329
Db 1465 GGGCTTACTGCTGACCTGCAAGGGCTGGGATCAGCTGGCTGGGGCTGCGCTCACTG 1524
QY 329 ----- 329
Db 1525 AGGAACTGGGCACTGAGTGGCCCTGATCCACATAAACACCCACTCTGCTGCTGCTGCT 1584
QY 329 ----- 329
Db 1585 ACGGTGCCCTGGGACCACTCTTTGGAGAACCCGACCAAGCTCTGCTCAGACTGCCAAC 1644
QY 329 ----- 329
Db 1645 CGGCAAGAGAGAGTGTGTGGGGAGGGGCTGGCTGCCACAGCTGTGCGCCGAGGG 1704
QY 329 ----- 329
Db 1705 CACTGTGGGGTCCAGGGGCCACCCAGTGTCTCACTGACAGCCAGTTCCTTGGGGCCAG 1764
QY 329 ----- 329
Db 1765 GAGTGCCTGAGAGATGCCAGTACTGCAAGGGGCTCCCAAGGAGTATGTGAATGCCAGG 1824
QY 329 ----- 329
Db 1825 CACTGTTTGGCGTCCACCCCTGAGTGTCAAGCCAGATGGCTCAGTACCTGTTTGA 1884
QY 329 ----- 329
Db 1885 CCGGAGGCTGACAGAGTGTGGCTGTGCCACTATGAAGACCTCCCTTCTGCGTGGCC 1944
QY 329 ----- 329
Db 1945 CGCTGCCCAAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2004
QY 329 ----- 329
Db 2005 GAGGAGGGCGCATGCCAGCTTGGCCCATCACTGACACCCACTCTCTGTGAGACTGGAT 2064
QY 329 ----- 329
Db 2065 GACAAAGGCTGCCCCCGGAGAGAGAGACAGCCCTCTGACGTCATCGTCTGCGGCTG 2124
QY 329 ----- 329
Db 2125 GTTGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2184
QY 329 ----- 329
Db 2185 CAGCAGAAGATCCGGAAGTACAGAGATGCGGAGACTGCTGCAAGAAAGAGAGCTGGTGGAG 2244

XX Proto-oncogene immunogen - used in vaccine for the prevention and
PT treatment of cancer
XX
PS Disclosure: Page 56-58; 81pp: English.
XX
CC This sequence represents the human HER2 cognate transgene (CTG).
CC Deletion of amino acids 1-731 of the encoded protein renders the CTG
CC non-transforming. HER2 is a tyrosine kinase-type receptor. This sequence
CC can be used in the cellular immunogen of the invention. The cellular
CC immunogen of the invention is for immunising against the product of a
CC target proto-oncogene, over-expression of which is associated with
CC cancer, comprises host cells transfected with a construct containing at
CC least one transgene related to the proto-oncogene and driven by a strong
CC promoter. The product of the transgene induces immunoreactivity to host
CC self-determinants on the product of proto-oncogene. The cellular
CC immunogens are used for protective vaccination against cancer (e.g.,
CC carcinoma of breast or colon, or various lymphomas) and for immunotherapy
CC of cancer. Use of the immunogen eliminates the need to isolate
CC immunogenic, HLA host-matched peptides. The method is not based on immune
CC recognition of a determinant defined by a cancer-specific mutation and
CC generates a systemic (anti-metastatic) response.
XX
SQ Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other:

Alignment Scores:
Pred. No.: 8.15e-98 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64.80% Gaps: 736
DB: 18 Gaps: 4

US-09-821-883-2 (1-690) x AAT71253 (1-4530)

OY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
DB 94 CGGCCCTCCAGCCGGGTCCAGCGGAGCATGGCGCGGAGCGAGTAGACCATG 153
OY 18 PheLeuPheLeuPhePhePhePhePheLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
DB 154 GAGCTGGCGGGCTGTGCGCGGTG-----GGGCTCCCTCGCGCCCTTCCGCC 204
OY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
DB 205 GGACCCCGGAGCACCAAGGTGTGCACCGGCACACATGAAGCTCGCGCTCCCTCCAGT 264
OY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
DB 265 CCGGAGACCCACCTGGACATGCTCCGCCACTCTACCAAGGCTCCAGGTGTGCAGGGA 324
OY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
DB 325 AACCTGGAACTCACCCTACCTGCGCCACCAATGCCAGCTGTCTTCGAGATATCAG 384
OY 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
DB 385 GAGGTGAGGGGTACGTGCTCATGCTCACAAACCAAGTGAAGGAGGTCCACCTGCAGAGG 444
OY 118 LeuArgTlleValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAsp 137
DB 445 CTGGGGATTGTGCGAGCACCCACCTCTTGGAGACAACTATGCGCTGGCGCTCTAGAC 504
OY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
DB 505 AATGGAAACCCGCTGACATACACACCCCTGTCCACAGGCGCTCCACAGAGGCTCTCGG 564
OY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsn 177
DB 565 GAGCTGACCTTGACCAACCTTCACAGAGATCTTGAAGAGAGGAGGTCTGATTCACAGGAAAC 624
OY 178 ProGluLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln 197

DB 625 CCCCAGCTTGTCTACAGGACACGATTTTGTGAGGACATCTTCACAAAGAACACAG 684
OY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
DB 685 CTGGCTCTACACATGATAGACACACACCGCTGTGGGCTGTCCACCCCTGTCTCCAGT 744
OY 218 CysLysGlySerArgCysTyrTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237
DB 745 TGTAAAGGCTCCCTCTCTGTGGGAGAGATTTGTAGATTTGTACAGACCTGACCGCAGT 804
OY 238 ValGluValGluGlyCysValAlaArgCysLysGlyProLeuProThrArgCysValGlu 257
DB 805 GTCTGTCCGGT 864
OY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerArgCysLeuAlaCysLeuHisPhe 277
DB 865 CAGT 924
OY 278 AsnHisSerGlyTleCysGluLeuHisCysProAlaLeuValThrTyrAspThrArgThr 297
DB 925 AACCAAGTGCATCTGTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 984
OY 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGluAlaSerCysValThrAla 317
DB 985 TTTCAGTCCATGCCCAATCCGAGGCGCGGTATTCATTCGGCGCCAGCTGTACTGTCC 1044
OY 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
DB 1045 TGTCCCTAACACTACCTTTTACAGAGACGTGGAGTCTGCACCTGCTGCTGCTGCTGCTGCT 1104
OY 329 ----- 329
DB 1105 AACCAAGAGTGACAGACAGAGATGAACACAGCGGTGTGAGAAGTGCACAGACCCCTGT 1164
OY 329 ----- 329
DB 1225 GCCAATATCCAGAGATTGT 1284
OY 329 ----- 329
DB 1285 GAGAGCTTTGATGGGACCCAGCCTCCACACTGCCCCCTCCAGCAGAGACACTCCAA 1344
OY 329 ----- 329
DB 1345 GTGTTGAGACTGTGAAGAGATCAGAGTTACTATATCATCTCAGATGTGCGGACAGC 1404
OY 329 ----- 329
DB 1405 CTGCCTGACCTCAGACGTCCTTCCAGAACCTGCAGAACTTAATCCGGGAGCAATTTCTGACA 1464
OY 329 ----- 329
DB 1465 GCGCCTACTCGCTGACCTGCAGAGGCTGGGATCAGCTGGGTGGCTGCTGCTGCTGCTGCT 1524
OY 329 ----- 329
DB 1525 AGGGAAGTGGGACAGTGGAGTGGCTTCATCCACATTAACACCCACCTGTGCTGTGCTGCT 1584
OY 329 ----- 329
DB 1585 ACGGTGCTTGGGACAGCTCTTTCGGAACCCGACACAGCTGTGCTGCTGCTGCTGCTGCT 1644
OY 329 ----- 329
DB 1645 CCGCAGAGAGACAGATGTGTGGGAGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704
OY 329 ----- 329
DB 1705 CACTGCTGGGTCCAGGCGCCACCCACCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764

QY 329 ----- 329
Db 1765 GAGTCCGTGGAGAAATCCGAGTACTGACAGGGGCTCCCCAGGAGATATGTGAATGCCAGG 1824
QY 329 ----- 329
Db 1825 CACTGTTTGGCGTGACACCTGAGTGTCAAGCCCCGAAATGAGTCACTGACCTGTTTGA 1884
QY 329 ----- 329
Db 1885 CCGAGGCTGACACAGTGTGTGGCTGTGCCACTATTAAGAACCTTCCCTTGTGCGTGGCC 1944
QY 329 ----- 329
Db 1945 CGCTGCCCAAGCGGTGAAACCTGACCTTCTACATGCCATCTGGAAGTTTCCAGAT 2004
QY 329 ----- 329
Db 2005 GAGAGGGCGCATGCGACGCTTGCCCATCAACTGACCACTCCTGTGTGGACCTGGAT 2064
QY 329 ----- 329
Db 2065 GACAAGGCTGCCCCCGGAGCAGAGACCAAGCCCTGTGACGTCCATGCTCTGCGGTG 2124
QY 329 ----- 329
Db 2125 GTTGGCATTTCTGCTGCTGTGCTCTTGCGGGGTGGTCTTTGGGATCTTCATCAAGCGACGG 2184
QY 329 ----- 329
Db 2185 CAGCAGAAATCCGGAAGTACACATGCGGAGACTGCTGCAAGAAAGGAGCTGGTAG 2244
QY 329 ----- 329
Db 2245 CCGCTGACACCTAGCGGAGCGATGCCCAACAGGCGAGATGCCGATCTGAAAGAGAG 2304
QY 329 ----- 329
Db 2305 GAGCTGAGGAAGTGAAGTGTGATCTGCGCTTTTGGCACAGTCTACAAGGGCATC 2364
QY 329 ----- 329
Db 2365 TGGATCCCTGATGGGAGATGTGAATAATTCAGTGCCATCAAGTGTGAGGGAAC 2424
QY 329 ----- 329
Db 2425 ACATCCCCCAAGCCAAAGAAATCTTAGAGCAAGCATACGTGATGGCTGTGGGC 2484
QY 329 ----- 329
Db 2485 TCCCCATATGTCTCCCGCTTCTGGGCATGTGCTGACATCATCAGCGTGCACGTGTGACA 2544
QY 329 ----- 329
Db 2545 CAGCTTATGCCCTATGCTGCTCTTAGACCATGTCCGGAAACCGCGGAGCCTGGGC 2604
QY 329 ----- 329
Db 2605 TCCAGAGACCTGCTGAAGTGTATGAGATTGCCAAGGGGATGAGTACCTGGAGAT 2664
QY 329 ----- 329
Db 2665 GTGGGGCTCGTACACAGGAGCTTGGCCGCTGGGAACGTGTGTCAAGTCCCAACAT 2724
QY 329 ----- 329
Db 2725 GTCAAAATTTACAGACTTGGGCTGGCTGGCTGCTGACATTGACAGACAGAGTAAACAT 2784
QY 329 ----- 329
Db 2785 GCAGATGGGGCAAGGTGCCCATCAAGTGAATGGCGCTGGAATTCATTCTCCGCGGCGG 2844

QY 329 ----- 329
Db 2845 TTCACCCACAGAGATGATGTGTGAGTTATGTGTGACTGTGTGGAGACTATGACTTTT 2904
QY 329 ----- 329
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCCGGAGATCCCTGACCTGTGAAAAGGGG 2964
QY 329 ----- 329
Db 2965 GAGGCGTCCCCAGCCGCCCATCTGCACCATTTGATGTCTACATGATCATGTGTAATGT 3024
QY 329 ----- 329
Db 3025 TGGATGATTGACTGTGAATGTGCGCAAGATTCGGGAGTGTGTCTGAATTTCCCGC 3084
QY 329 ----- 329
Db 3085 ATGCCAGGAGACCCACGCTTTGTGTCATCCAGATGAGACTTGGCCACAGCAT 3144
QY 329 ----- 329
Db 3145 CCGTTGACAGACCTTCTACCGCTCACTGCTGAGAGACATGACATGGGGACCTGGT 3204
QY 330 ----- 330
Db 3205 GATGCTGAGAGATATCTGTGACCCAGAGGCTTCTTGTCCAGACCTGCCCCGGGC 3264
QY 331 AlaGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyIleVal 350
Db 3265 GCTGGGGCATGTGTCACACAGGACCGACCGACATCATCTACAGAGATGGCGGTGGGAC 3324
QY 331 LeuThrLeuGlyLeuGluProSerGluGluIleAlaProAlaGlySerProLeuAlaProSer 370
Db 3325 CTGACACTAGGGCTGAGCCCTCTGTAAGAGAGGGCCCCAGGTCTCCACTGGCACCTCC 3384
QY 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlySerIleu 390
Db 3385 GAAGGGCTGGCTCCGATGTATTGTATGTGATGCTGACCTGGGAATGGGGACCCAAAGGGCTG 3444
QY 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410
Db 3445 CAAAGCTCCCCACACATGAGACCCCGCTCTACAGCGGTACAGAGAGAGACCCACACATA 3504
QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrGlySerProGluProIle 430
Db 3505 CCGCTGCCCTGTGAGACTGATGTGCTACGTGGCCCCCTGACCTGACGCCACCCAGCCTGAA 3564
QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 450
Db 3565 TATGTGAACACAGCCAGATGTGTGGCCAGCCCTTGGCCCGAGAGGGCCCTTGCT 3624
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaValGlyIleAsnProGluTyrLeuThr 470
Db 3625 GCTGCCACACTGCTGTGGCCACTGTGAAAGGGCCCAAGACTCTTCCCCAGGGAAGAT 3684
QY 471 GlyValValIleAspValPheAlaPheGlyGlyAlaValGlyIleAsnProGluTyrLeuThr 490
Db 3685 GGGGTCTGCAAAAGACTTTTGGCTTGGGGGTGCGTGGGAACCCGAGACTGTGACA 3744
QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 510
Db 3745 CCGCAGGAGAGAGCTGCCCTCAGCCACCTCTCTGCTTCCAGCCACCTTGCAG 3804
QY 511 AsnLeuTyrTyrPhePheGlnAspProProGluArgGlyAlaProProSerThrPheIys 530
Db 3805 AACCTCTATTACTGGGACAGAGACCCAGAGGCGGGGGCTCCACCCAGCACCTTCAAA 3864
QY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaIlePro 550
Db 3865 GGGACACTAGGACAGAAACCCAGATACCTGGGTGTGGAGCTGCA -----GTGTGA 3918
QY 551 AlaArgSerProSerPro 556

329 329
Db 1285 GAGAGTTTGATGGGAGCCAGCCTCCAACTGCCCCCTCCAGCCAGAGCACTCCAA 1344
329 329
Db 1345 GTGTTGAGACTGTGAGAGATACAGAGTTACTATACATCATGACATGCGCGAGACG 1404
329 329
Db 1405 CTGCTGACCTGACGCTCTTCCAGAACTGCAAGTAATCCGGGAGCAATTCGCAAAAT 1464
329 329
Db 1465 GGGCGCTACTCGCTGACCTGCAGAGGCTGGGCAATCAGCTGGGCTGGGCTGACTG 1524
329 329
Db 1525 AAGGAACCTGGGAGTGGAGCTGGCCCTCATCCACATACACCCACTGCTTCGTGAC 1584
329 329
Db 1585 ACGGTGCCCTGGAGACCACTCTTTGGGAACCGCACCAAGCTCTGCTCCACACTGCCAAT 1644
329 329
Db 1645 CGGCCAGAGAGAGAGTGTGGGGCGAGGGCTGGCCCTGCACCAAGCTGTGCCCGAGGG 1704
329 329
Db 1705 CACTGTGGGGTCCAGAGGCCACCAGTGTGTCACAGCAGAGTCTCTGGGGGCGAG 1764
329 329
Db 1765 GAGTGGCTGAGAGAAATGCCGAGTACTGACAGGGCTCCCAAGGAGTATGTGAATGCCAG 1824
329 329
Db 1825 CACTGTTGCCGTGCCACCTGAGTGTCAAGCCCAAAATGGCTGACTGACTGTTTGA 1884
329 329
Db 1885 CGGAGGCTGACCAAGTGTGTGGCTGTGCCACTATAGAGACCTCCCTTCCGTGGCC 1944
329 329
Db 1945 CGCTGCCCAAGCGGTGTGAACCTGACTCTCTCTACATGCCCATCTGGAATTTCCAGAT 2004
329 329
Db 2005 GAGAGAGGGGCAATGCCAGCCTTGCCCCATCACTGCACCCACTCCTGTGGACCTGAT 2064
329 329
Db 2065 GACAAAGGCTGCCCCCGCAGAGAGAGCCAGCCCTGTGACGTCCATCTCTGCGGTG 2124
329 329
Db 2125 GTTGGCATTTCTGCTGCTGTGCTTTGGGGTGGTCTTTGGGATCCTCATCAAGCAGCG 2184
329 329
Db 2185 CAGCAGAAGATCCGGAAGTACAGATGGGAGACTCTGCAGAGAAACGAGCTGTGGAG 2244
329 329
Db 2245 CCGCTGACACTAGCGGAGCGATGCCAAACAGAGCGCAGATGGGATCCTGAAAGAGACG 2304
329 329
Db 2305 GAGCTGAGGAAGGTGAAGGTCTTGGATCTGGCGCTTTTGGCACAGTCTCAAGAGGCAATC 2364

329 329
Db 2365 TGGATCCCTGATGGGAGAAATGTGAAAATTCAGATGGCCATCAAAAGTGTGAGGGAAC 2424
329 329
Db 2425 ACATCCCCCAAGCCAAAGAAATCTTAGACAGCATAGTGTGCTGTGGCG 2484
329 329
Db 2485 TCCCATATGTCTCCCGCTTCTGGGCATCTGCTGACATCCACGGTGCACCTGTTGACA 2544
329 329
Db 2545 CAGCTTATGCCCTATGGCTGCCCTTAGACATGTCCGGGAAAAACCGCGAGCCTGGGC 2604
329 329
Db 2605 TCCAGAGACCTGCTGAATGTGTATGACAGATTGCCAAGGGATGACTACCTGAGAGAT 2664
329 329
Db 2665 GTGGGCTGTACACAGGACTTGGCCGCTCGGACGTCTGTGTCAGAGTCCCAACAT 2724
329 329
Db 2725 GTCAAAATTAAGACTTGGGGCTGGCTGGCTGTGACATTGACGAGACAGATACAT 2784
329 329
Db 2785 GAGATGGGGGCAAGGTGCCATCAAGTGAATGGCGCTGAGTCCATCTCCGCGCGG 2844
329 329
Db 2845 TTCACCCACAGAGTATGTGGAATTATGTTGACTGTGTGGAGCTGATGACTTT 2904
329 329
Db 2905 GGGGCCAAACCTTAGATGGATCCACGCCGGAGATCCGTGACTGTGGAAGAGGG 2964
329 329
Db 2965 GAGCGGCTCCCGAGCCCACTGACCATTTGATGTACATGATCATGTCAATGT 3024
329 329
Db 3025 TGAATGTTGACTCTGAATGTGGCCAAATTCGGGAGTTGTGTCTGAATTTCCCGC 3084
329 329
Db 3085 ATGGCCAGGAGACCCCGAGCGTTTGTGTATCAGATGAGACTTGGGCCAGCCAGT 3144
329 329
Db 3145 CCGTTGAGAGACCTTACCGCTCACTGTGAGAGAGATGACATGGGGACCTGATG 3204
329 329
Db 3205 GATGCTGAGAGATCTGTGATCCCGCAGAGGCTTTCTGTCCAAACCTGCCCGCGGGC 3264
329 329
Db 3311 AAGlyGlymeValHisHisArgHisArgSerSerThrArgSerGlyGlyAsp 3350
329 329
Db 3325 GCTGGGGGATGATGTCACCAAGGACCCAGCTCATCAACGAGATGGCGGTGGGAG 3324
329 329
Db 3351 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 370
329 329
Db 3325 CTGACACTGAGGGCTGAGCCCTCTGAAGAGAGGAGGCCCCAGAGCTTCACCTGGACCTCC 3384
329 329
Db 3371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaValysGlyLeu 390
329 329
Db 3385 GAGGGGCTGGCTCCGATATATTGATGTGACTGGAAATGGGGCAGCCAAAGGGGCTG 3444
329 329
Db 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410

Oy	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisproCysSerProMet	217
Dp	665	CTGGCTCTCAACAGTATAGAACCAACCGCTCTGGGGCTGGCACCCCTGTTCTCCGATG	744
Oy	218	CysAlaGlySerArgCysTrpArgLysIuSerSerGluAspCysGlnIserLeuThrArgThr	237
Dp	745	TGTAAAGGGCTCCGGCTCTGGGGAGAGATTCTGAGAGATTGTTCAGAGCCTGACGGGCACT	804
Oy	238	ValCysAlaGlyLysCysAlaIaArgCysLysGlyProLeuProThrAspCysHisGlu	257
Dp	805	GTCGTGCTCCGGTGGCTGTGGCCCTGGCAAGGGGGCACTGCCACATGACTGTCGCATGAG	864
Oy	258	GlnCysAlaAlaGlyCysThrArgLysProLysHisSerAspCysLeuAlaLysLeuHisPro	277
Dp	865	CAGTGTGCTGCCGGCTGACAGGGGCCCAACACATGTGACTGCTGGCTGCTCCATTC	924
Oy	278	AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrCysThrAspThr	297
Dp	925	AACCAACAGTGGCAATCTGTAGAGCTGCATGCCGCCCTGTCACTACCAACACAGACAGC	984
Oy	298	PheGluSerMetProAsnProGlnGlyArgTyrThrPheGluValAserCysValThrAla	317
Dp	985	TTTGAGTCCATGCCCAATCCCGAGGGGGGTATATACATTCGGGGCCAGCTGTGACTGCC	1044
Oy	318	CysProTyrAsnTyrLeuSerThrAspValGlySer-----	329
Dp	1045	TGTGCTTACAACTACTCTTTTACGGAGAGTGGAGTCTCGACACCCTGCTGCCCCCTGCAC	1104
Oy	329	-----	329
Dp	1105	AACCAAGAGTGCACAGCAGAGATGGAAACACAGCGGTGTGAGAAAGTGCAGAAAGCCCTGT	1164
Oy	329	-----	329
Dp	1165	GCCGAGAGTGTCTATGTGTCTGGGCAATGGAGCACTTGGAGAGGTGAGGGCACTTACCACT	1222
Oy	329	-----	329
Dp	1225	GCCAAATATCCAGAGAGTGTGTGGCTGCAGAAAGATCTTTGGAGAGCTTGCAATTCGCGC	1284
Oy	329	-----	329
Dp	1285	GAGAGCTTTGATGGGGACCAAGCCTCCAAACATCTGCCCCGCTCCAGCCAGAGCACTCCAA	1344
Oy	329	-----	329
Dp	1345	GTTGTTTGAGACTCTGGAAAGATTCACAGTTTACTTATACATCTCACACATGGCCGGACAGC	1404
Oy	329	-----	329
Dp	1405	CTGCGTGAACCTCAGCGTCTTTCAGAAACCTGCAGATTAATCCGGGAGAGAAATTCGACAAAT	1464
Oy	329	-----	329
Dp	1465	GGCGCCTAATCTGCTGACCTTGCAAGGGCTGGGAGTCAGCTGGCTGGGGCTGGCCTCACTG	1524
Oy	329	-----	329
Dp	1525	AGGGAAGTGGGCACTGGAGACTGGCCCTCATCCACATTAACACCCACACTCTGCTTGGTGAC	1584
Oy	329	-----	329
Dp	1585	ACGGTGCCCTGGAGACAGCTCTTTCCGAAACCGCGACCAACAGCTGTGCTCCACACTGCAAC	1644
Oy	329	-----	329
Dp	1645	CGGCGAGAGAGCAGTGTGTGGGGAGAGGCTGGCTGCCACCAAGCTGTGCGCCGAGGG	1704
Oy	329	-----	329
Dp	1705	CACGTCTGGGGTCCAGGGGCCCAACCACTGTGTCAACTGCAGCCAGTTCTTTCGGGGCCAG	1764
Oy	329	-----	329

Db	1765	GAGTCCGTGGAGGAATCCCAAGTACTGACAGGGGCTCCCAAGGAGATGTGAATGCCAGG	1824
QY	329	-----	329
Db	1825	CACGTGTTTGCCGTCCACCCCTGAGTGTCAAGCCCCAGAAATGGCTCACTGTGTTTGA	1884
QY	329	-----	329
Db	1885	CCGGAGGCTGACCAAGTGTGTGGCCTGTGCCACTAATAAGACCCTCCCTTCTGCGTGGCC	1944
QY	329	-----	329
Db	1945	CGCTGCCCCAGCGGTGTGAACCTGACCTCTCCATCATGATGCCACTGTGAAGTTCCAGAT	2004
QY	329	-----	329
Db	2005	GAGGAGGGCGCATCCAGCCTTGCCCATCACTGACACCCTCTGTGTGACCTGGAT	2064
QY	329	-----	329
Db	2065	GACAAGGCTGCCCCCGGAGCAGAGAGGCGAGCCCTGTGACGTCCATGCTCTGCGGTG	2124
QY	329	-----	329
Db	2125	GTTGGCATTTGCTGTGTGTGTCGTCGTCCTTGGGGGTGGTCTTTGGGATCCTCATCAAGCAGGG	2184
QY	329	-----	329
Db	2185	CAGCAGAAGATCCGGAAGTACACAGTCGAGACCTGCTGACGAAGAACGAGCTGTGGAG	2244
QY	329	-----	329
Db	2245	CCGCTGACACCTAGCGGAGGAGTGGCCCAACAGCGGAGATGGCGATCTCTGAAGAGAGG	2304
QY	329	-----	329
Db	2305	GAGCTAGAGAGGTGAAGTGCTTGATCTGGCGCTTTTGGCACAGTCTACAAGGCGATC	2364
QY	329	-----	329
Db	2365	TGGATCCCTGATGGGGAGAAATGTGAATAATCCAGTGGCCATCAAAATGTTGAGGGAAAC	2424
QY	329	-----	329
Db	2425	ACATCCCCCAAGGCCAACAAAGAAATCTTAGACGAAGATACGTGATGCTGGTGGGC	2484
QY	329	-----	329
Db	2485	TGCCCATATGTCTCCCGCTTCTGGGCATCTGCCCTGACATCCACGGTGCAGCTGGTGACA	2544
QY	329	-----	329
Db	2545	CAGCTATGCCCTATGGGTGCTCTTAGACCATGTCCGGGAAACCGCGAGCGCTGGGC	2604
QY	329	-----	329
Db	2605	TCCAGAGACTGTGAACCTGGTGTATGACAGATTGCCAAGGAGTGAAGTACCTGGAGAT	2664
QY	329	-----	329
Db	2665	GTTGGCGCTGTACACAGAGGACTTGGCCGCTCGGAAGCTGTGTCAAGATGCCAACAAT	2724
QY	329	-----	329
Db	2725	GTCAAATTTACAGACTCGGGCTGGCTCGGCTGCTGGACATTGACGAGACAGATACCAT	2784
QY	329	-----	329
Db	2785	GCAATGGGGGCAAGGTGCCATCAAGTGGATGGCGCTGAGATCCATTCTCCGCGGGGG	2844
QY	329	-----	329

```

Db 2845 TTCACCCACGAGATGATGTGTGAGATTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904
OY 329 ----- 329
Db 2905 GGGCCAAACCTTAGATGGATGCCAGCCGGGAGATCCCTGACCTGTGGAAAAGGG 2964
OY 329 ----- 329
Db 2965 GAGGGGCTGCCAGCCGCCCATCTGCACATGTATGTCTACATGATCATGTCAAAATGT 3024
OY 329 ----- 329
Db 3025 TGGATGATGACTGTGAATGTGGCCCAAGATTCCGGGAGTTGTGTGTAATTTCCCGC 3084
OY 329 ----- 329
Db 3085 ATGGCCAGGAGACCCCAAGCGTTTGTGTATCCAGATGAGACTTGGGCCAGCCAGT 3144
OY 329 ----- 329
Db 3145 CCCTTGGACAGACCTTCTACCCCTCAGTGTGAGAGCAGATGACATGGGGAGCTGGTG 3204
OY 330 -----Gly 330
Db 3205 GATCCTGAGAGATATCTGTACCCAGAGGGCTTCTTCTGTCCAGACCTGCCCGGGGC 3264
OY 331 AlAGlYlmeTValhshshsArgHisArgSerSerThraYserGlyGlyYasp 350
Db 3265 GCTGGGGGCAATGTCCACACAGGACCGCAGCTCATCTACAGAGAGTGGGGTGGGAC 3324
OY 351 LeuthreuglYleuGluprosersgluGlulAlaproasgersproleuAlaproser 370
Db 3325 CTGACACTAGGGCTGAGACCTCTGAAGAGAGGAGCCGCCAGGTCTCCAGTGGACCTCC 3384
OY 371 GluGlYlAGlYserAspValhshshsArgHisArgSerSerThraYserGlyGlyYasp 390
Db 3385 GAAAGGGCTGCTCGATGTATTGATGTGATGCTGGGAATGGGGGACGCCAAGGGGCTG 3444
OY 391 GlinserleuprothrhshsAspProserProleuGlhArgYrSerGlYaspProthVal 410
Db 3445 CAAAGCCTCCACACATGACCCACGAGCCCTCTACAGCGCTGACGTAGAGAGCCACAGTA 3504
OY 411 ProleuprosersgluThraYspGlyTyrValAlaproleuthrYcysSerProglu 430
Db 3505 CCCCTGCCCTTGAGACTGATGGTACGTTGCCCTGACCTCAGCCGCCACGCTGAA 3564
OY 431 TyrValshnglnProasPvalhArgProgluProserProasYrgluGlYleuPro 450
Db 3565 TATGTGAACGACGAGATGTTGGGCCAGCCCTTCGCCGAGAGGGGCCCTGCTCCT 3624
OY 451 AlAlAlarproAlaGlYlAlaThleuGlulArgAlaLysThrlLeuSeiProglYLysAsn 470
Db 3625 GCTGCCCGACCTGCTGTGGCCACCTGTGAAAGGGCCAAAGCTCTCTCCCGAGGAAAGAT 3684
OY 471 GlYAlaYlAlYsAsPValhPheAlaPheGlYGlYAlaValGlYAsnProgluTyrLeuThr 490
Db 3685 GGGGTCTGTCAAGACGTTTTTGGCTTTGGGGGTCCCTGGGAGAACCCCGAGTACTTACA 3744
OY 491 ProGlInglYlAlaAlaProlGlnProhshsProProAlaPheSerProAlaPheAsp 510
Db 3745 CCCAGGAGAGAGCTGCCCTCAGCCGCCACCTCTCTGCTTCACGCCAGCTTGCAC 3804
OY 511 AsnleuTyrTyrTrpAspGlnAsPProProgluArgGlYAlaProlProserThrPheLys 530
Db 3805 AACCTCTATTACTGTGGACCGAGCCACAGAGCGGGGCTCCACACGACACTTCAAA 3864
OY 531 GlYThrProThrhshshsAsnProgluTyrLeuGlYleuAspValhProlAlaAlaLysPro 550
Db 3865 GGGACACTTACGGGAGAAACCGAGTACTGTGGTGTGACGTGCA-----GTGTGA 3918
OY 551 AlArYserProserPro 556
Db 3919 ACCAGAGGCCAACTCG 3936

```

```

RESULT 9
ABN85585
ID ABN85585 standard; DNA: 4530 BP.
XX
AC ABN85585;
XX
DT 09-SEP-2002 (first entry)
XX
DE Human HER2-neu SEQ ID NO 11.
XX
KW Human; EGFR; HER2-neu; chemotherapeutic regimen; tumour; cancer;
KW receptor tyrosine kinase; epidermal growth factor receptor;
KW gene expression; ds.
XX
OS Homo sapiens.
XX
PN W020024413-A2.
XX
PD 06-JUN-2002.
XX
PF 09-NOV-2001; 2001WO-US43035.
XX
PR 01-DEC-2000; 2000US-250122P.
PR 04-DEC-2000; 2000US-250469P.
PR 11-JUN-2001; 2001US-0871177.
XX
PA (RESP-) RESPONSE GENETICS INC.
XX
PI Danenberg KD.
XX
DR WPI: 2002-537460/57.
XX
PT Determining chemotherapeutic regimen of receptor tyrosine kinase
PT targeted agent for treating tumor by examining EGFR and/or HER2-neu
PT mRNA amount in tumor cells, comparing it to predetermined threshold
PT expression level.
XX
PS Disclosure; Page 124-125; 125pp; English.
XX
CC The invention relates to determining the chemotherapeutic regimen of
CC receptor tyrosine kinase targeted agent for treating tumour by amplifying
CC mRNA from tumour and non-malignant tissues using a primer pair that
CC hybridises to epidermal growth factor receptor (EGFR) and/or HER2-neu
CC gene (I), quantitating and obtaining differential expression levels of
CC amplified mRNA and comparing the differential expression levels and
CC threshold levels for expression of (I). The method is useful for
CC assessment of clinical treatment of a patient and as a diagnostic or
CC prognostic tool for a range of cancers including breast, head and neck,
CC lung, oesophageal and colorectal cancer. The present sequence is that of
CC the human HER2-neu DNA sequence used in methods of the invention.
XX
SQ Sequence 4530 BP; 922 A; 1382 C; 1346 G; 880 T; 0 other:
XX
Alignment Scores:
Pred. No.: 8,15e-98 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64,80% Indels: 736
DB: 24 Gaps: 4
US-09-821-883-2 (1-690) x ABN85585 (1-4530)
OY 2 ArgAlaAlaProleuLeuAlaAlaAla-----AlaserLeuSerLeuGly 17
Db 94 CGCGCCCTCCAGCGGGGTCCAGCGAGCATGGGCGCGAGCGCAGTACAGCACCATG 153
OY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValAlaAlaLysGluLeuAlaArg 37
Db 154 GAGCTGGCGGCGCTGTGGCCGCTG-----GGGCTCTCTCCCTCGCCCTTTCCCGCC 204
OY 38 GlYAlaAlaSerThrhGlnValCysThrhGlyThraSpMetLysLeuArgLeuProAlaSer 57

```

Db 205 GGAGCCGCGACACCACGAGTGTGCACCGCACAGACATGAAGCTGCGCTCCCTCCAGT 264
 Oy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
 Db 265 CCCGAGACCCACCTGGACATGTCTCCGACCTACACAGGGCTGCAGAGTGGTGACGGA 324
 Oy 78 AsnLeuGlnLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
 Db 325 AACCTGGAACTCACCCTACCTGCCCCAACCAATCCAGCCCTGCTCTCTCAGAGATTCAG 384
 Oy 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
 Db 385 GAGGTGACGGGCTACGTCTCATGCTCACAACCAAGTAGAGGAGCTCCCACTGCAGAG 444
 Oy 118 LeuArgIleValArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAsp 137
 Db 445 CTGGCGATTGTGCAGGACCCAGCTCTTGTAGGCAACTATGCCCTGCGCTGTAGAC 504
 Oy 138 AsnGlyAspProLeuAsnAsnThrProValThrGlyAlaSerProGlyGlyLeuArg 157
 Db 505 AATGAGACCCGCTGACAAATACCAACCCCTGTACAGGGGCTCCCGAGAGGCTTGCGG 564
 Oy 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuGlyGlyValLeuIleGlnArgAsn 177
 Db 565 GAGCTGACAGCTTCGAAAGCTCAGAGATCTTGAAGAGAGGGCTCTGATCCAGCGAAC 624
 Oy 178 ProGlnLeuCysTyrGlnAspThrIleLeuTyrPlysAspIlePheHisLysAsnAsnGln 197
 Db 625 CCCCAGCTCTCTACACAGACACGATTTTGTGGAAGGACATCTCCACAGAACACACAG 684
 Oy 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
 Db 685 CTGGCTCTCACACTGATGACACCAACCGCTCTGCGGCTGCGACCCCTGTTCTCCGATG 744
 Oy 218 CysLysGlySerArgCysTyrP6LysLeuSerGlnAspCysGlnSerLeuThrArgThr 237
 Db 745 TGTAAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGATGTCAGAGCGCTGACGCGCACT 804
 Oy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisLysGlu 257
 Db 805 GTCTGTGCGGGTGTGTGCGCGCTGCAAGGGCCACTGCCCCACTGACTGCTGCGCATAG 864
 Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
 Db 865 CAGTGTGCTGCGGCTGCACGGGCCCAAGCACTGTGACTGCTGCTGCTGCTCCACTTC 924
 Oy 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
 Db 925 AACCACTGGCAATCTGTGACTGCACTGCCAGCCCTGTGCTACCTACACACACAGACAG 984
 Oy 298 PheGluSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
 Db 985 TTTGAGTCCATGCCCAATCCGAGGGCGGTATACATTGCGCGCAGCTGTGTGACTGCC 1044
 Oy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer 329
 Db 1045 TGTCCCTCAACTACCTTTCTACGAGAGCTGGGATCTCAGCCCTGCTGCTGCCCTCGAC 1104
 Oy 329 329
 Db 1105 AACCAAGAGGTGACAGCAGAGAGATGAAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGT 1164
 Oy 329 329
 Db 1165 GCCCGAGTGTCTATGTCTGGGACATGGAGCACTTGCAGAGAGTGAGGGACGTTACCACT 1224
 Oy 329 329
 Db 1225 GCCAATATCAAGAGATTGTGCTGCTGCAAGAAGATCTTTGGAGGCTGCAATTTCTGCGG 1284
 Oy 329 329

Db 1285 GAGAGCTTTGATGGGGACCCAGCCTCCAAACACTGCCCCGCTCCAGCCAGACAACTCCAA 1344
 Oy 329 329
 Db 1345 GTGTTGAGACTCTGGAAGAGATCAGAGTTACCTATACATCTACGATGGCCGACAGC 1404
 Oy 329 329
 Db 1405 CTGCCGTGACCTCAGGCTTCTCCAGAACCTGCAGATATCCGGGGAGCAATTCTGCACAA 1464
 Oy 329 329
 Db 1465 GGGCCTACTGCTGACCTCGCAAGGCTGGGATCAGCTGGCTGGGCTGCGCTCACTG 1524
 Oy 329 329
 Db 1525 AGGCACTGGGCACTGAGCTGGCCCTCATTCACACATAACACCACCTCTGCTTGCTGAC 1584
 Oy 329 329
 Db 1585 ACGTGCCCTGGGACCACTCTTTCGGAACCCGCAACAGCTCTGCTCAGACTGCCAAC 1644
 Oy 329 329
 Db 1645 CGGCCAGAGACAGATGTGTGGCGAGGGCTGGCCCTGCACACAGCTGTGCGCCGAGG 1704
 Oy 329 329
 Db 1705 CACTGCTGGGTTCCAGGGCCACCCAGTGTGTCACTGACCCAGTTCCTTGGGGCCAG 1764
 Oy 329 329
 Db 1765 GAGTCCGTGAGAGAAATGCCAGTACTGTCAGGGGCTCCCGAGGAGTATGTGAATGCCAG 1824
 Oy 329 329
 Db 1825 CACTGTTTGCCGTGCCACCTGAGTGTGACGCCAGAAATGGCTCAGTACCTGTTTGA 1884
 Oy 329 329
 Db 1885 CCGAGGCTGACCACTGTGTGCGCTGTGCCACTATTAAGAACCTCCCTTGTGCTGCGC 1944
 Oy 329 329
 Db 1945 CGTGCCCCCAGGCGGTGTAACCTGACCTCTCTACATGCCCATGTGGAATTCAGAT 2004
 Oy 329 329
 Db 2005 GAGAGGCGCATGCCAGCTTGTGCCCATCAACTGCACCACTCCTGTGTGAGCTGAT 2064
 Oy 329 329
 Db 2065 GACAAGGCTGCCCGCCGAGCAGAGACCAAGCCCTGTGAGTCCATGCTGTGCGGTG 2124
 Oy 329 329
 Db 2125 GTTGGCATTCTGT 2184
 Oy 329 329
 Db 2185 CAGCAGAAGATCCGGAAGTACACGATGCGAGACTGTGCAAGAAACGAGAGCTGTGAG 2244
 Oy 329 329
 Db 2245 CCGCTGACACCTAGCGGAGCGATGCGCAACCAAGCGCAGATGCGGATCTCTAAGAGAGC 2304
 Oy 329 329
 Db 2305 GAGCTGAGAAAGTGAAGTGTGCTTGATCTGCGCTTTTGGCAGAGTCTACAAGGCGATC 2364
 Oy 329 329
 Db 2365 TGGATCCCTGATGGGAGAAATGTGAAAATTCAGTGGCCATCAAGTGTGAGGGAAGAAC 2424

QY 329 ----- 329
 Db 2425 ACATCCCCAAGCAACAAAGAAATCTTAGACGAGCATACGTGATGCTGTGTGGC 2484
 QY 329 ----- 329
 Db 2485 TCCCCATATGTCTCCGCCCTTGTGGGATCTGCTGACATCCAGGTGACGTGTGACA 2544
 QY 329 ----- 329
 Db 2545 CAGTTATGCCCTATGGCTGCTCTTAGACCATGTCCGGGAAAAACCGGGAGCGCTGGGC 2604
 QY 329 ----- 329
 Db 2605 TCCGAGACCTGTCTGAACGTGGTATGACAGATTCGAAGGGATAGCTACCTGGAGAT 2664
 QY 329 ----- 329
 Db 2665 GTGGGCTCGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGTCAAGATCCCAACCAT 2724
 QY 329 ----- 329
 Db 2725 GTCAAAATTTACAGACTTGGGCTGGCTGGCTGGACATTGACGAGACAGATACCAT 2784
 QY 329 ----- 329
 Db 2785 GCAGATGGGGGCAAGTGTCCCATCAAGTGAATGGCGCTGAGTCCATTTCCGCCGGC 2844
 QY 329 ----- 329
 Db 2845 TTCAACCAACAGAGATGATGTGTGATTTAGTGTGACTGTGTGGAGCTGATCACTTT 2904
 QY 329 ----- 329
 Db 2905 GGGCCAAACCTTACGATGGATGCCAGCCCGGGAGATCCCTGACCTGTGGAAAAAGGG 2964
 QY 329 ----- 329
 Db 2965 GAGCGCTGCCCAAGCCGCCCATCTGCACATGTATGTCTACATGATCATGTCATAATGT 3024
 QY 329 ----- 329
 Db 3025 TGGATGATGACTGTGATGTGCGCCCAAGATTCGGGAGTTGGTGTCAATTTCCCGC 3084
 QY 329 ----- 329
 Db 3085 ATGCCAAGGAGCCCAAGCGCTTGTGTGTCATCCAGATGAGACTTGGGCCACGCACT 3144
 QY 329 ----- 329
 Db 3145 CCCTTGAGACAGACCTTCTACCGCTCACTGTGAGAGACATGACATGGGGAGCTGGTG 3204
 QY 330 ----- 330
 Db 3205 GATGCTGAGAGATGTCGTGATCCCAAGGCTTCTTCTGTCCAGACCTGCCCGGGC 3264
 QY 331 AAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3324
 Db 3265 GCTGGGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3384
 QY 351 LeuThrLeuGlyLeuGluProSerGluGluGluGluGluGluGluGluGluGluGluGlu 370
 Db 3325 CTGACACTAGGGCTGAGACCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3384
 QY 371 GAG 390
 Db 3385 GAG 3444
 QY 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410
 Db 3445 CAAAGCTCCCAACATGACCCCAAGCCCTCTACAGCGGATGAGAGAGAGAGAGAGAG 3504

QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
 Db 3505 CCCCCCTGCTGTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3564
 QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuPro 450
 Db 3565 TATGTGAACCAAGCAGATGTTGGGCCCAAGCCCTTGTGGCCCAAGAGAGAGAGAGAG 3624
 QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluValArgAlaThrLeuSerProGlyLysAsn 470
 Db 3625 GCTGCCGAGACTGCTGTGGTCCCACTGTGAAAGGGCCAAAGACTCTCTCCCAAGGAAAT 3684
 QY 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
 Db 3685 GGGGTGCGCAAAAGAGCTTTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 3744
 QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
 Db 3745 CCCCCAGGAGAGCTGCCCCCTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3804
 QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
 Db 3805 AACCTATATCTGTGGACACAGACCAACAGAGGGGGGCTCCACACGACACTTCAAA 3864
 QY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaPro 550
 Db 3865 GGGACACCTACGCGAGAGAACCCAGAGTACTGGGTCTGACGTGCCA-----GTGTGA 3918
 QY 551 AlaArgSerProSerPro 556
 Db 3919 ACCAGAAAGCCAACTCCG 3936
 RESULT 10
 ABK83918
 ID ABK83918 standard; cDNA; 4530 BP.
 AC ABK83918;
 XX
 DT 14-AUG-2002 (first entry)
 DE Human cDNA differentially expressed in granulocytic cells #489.
 XX
 KW Human; ss; granulocytic cell; DNA chip; bacterial infection;
 KW viral infection; parasitic infection; protozoal infection;
 KW fungal infection; sterile inflammatory disease; psoriasis;
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
 KW adult respiratory distress syndrome; inflammatory bowel disease;
 KW Crohn's disease; ulcerative colitis; periodontal disease;
 KW granulocyte activation; chronic inflammation; allergy.
 XX
 OS Homo sapiens.
 XX
 PN W0200228999-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 03-OCT-2001; 2001WO-US30821.
 XX
 PR 03-OCT-2000; 2000US-237189P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
 XX
 DR WPI; 2002-435328/46.
 XX
 PT Detecting granulocyte activation by detecting differential expression
 PT of genes associated with granulocyte activation, which serves as
 PT diagnostic markers that is useful for monitoring disease states and
 PT drug toxicity
 XX
 PS Claim 1; SEQ ID NO 489; 114pp; English.

XX The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysts as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease; by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA, M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Alignment Scores:	
Pred. No.:	8.15e-98
Score:	2423.00
Percent Similarity:	41.14%
Best Local Similarity:	40.75%
Query Match:	64.80%
DB:	24
Length:	4530
Matches:	524
Conservative:	21
Mismatches:	5
Indels:	736
Gaps:	4

US-09-821-883-2 (1-690) x ABK83918 (1-4530)

OY	2	ArgAlaAlaProLeuLeuAlaValgla-----AlaSerSerLeuGly	17
		: :	
Db	94	CGCGCCCTCCAGCGGGGTCCAGCCGGAGCCATGGGGCCGAGCCCGACGTAGACACCATG	15
OY	18	PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaValGluLeuAlaArg	37
		:	
Db	154	GAGCTGGGGCGGCTGTGGCGGTGG-----GGGTCTCTCTCGGCTCTTGCCTCC	204
OY	38	GlyAlaAlaSerThrGlnValAlcysThrGlyThrAspMetLsleArgLysProAlaSer	57
Db	205	GGAGCCGGGAGACCCCAAGTGTGGCACCGGCACAGACATGAAGCTGGGCTCTCTGCCAGT	264
OY	58	ProGluThrHisLeuAspMetLeuArgHisLeuTyrgingLysGlnValValGlnGly	77
Db	265	CCCGAGACCACCTGGACATGCTCCGCCACCTTACACAGGGCTGGCAGGTGTCAGGGA	324
OY	78	AsnLeuGluLeuThrTyrgLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	325	AACCTGGAACCTACCTACCTGCCACCAATGCACACCTCTCTCTCTCAGAGATATCCAG	384
OY	98	GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
Db	385	GAGGTGCAGGGCTACGTGCTCATCTCCACAAACCAAGTAGAGCAGGTCCACTCAGAGG	444

QY	118	LeuArg11LeuValArgGlyThrGlnLeuPheGluIAspAsnTyrAlaLeuAlaValLeuAsp	137
-Db	445	CTGGGATTTGGCCAGGACCACAGCTCTTTGAGAGCAACATATGACCCTGGCCCTGCTATAC	504
QY	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
Db	505	AATGAGACCCCGCTGAAACATATACCACCCCTGTACAGGGGGCCTCCCGAGAGCCCTGGG	564
QY	158	GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn	177
Db	565	GAGGTGACGCTTCACAAACCTTCACAGAGATCTTGAAAGAGGGGTCTTGATTCAGCGGAAAC	624
QY	178	ProGlnLeuGlyTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGln	197
Db	635	CCCCAGCTCTGCTACCGAGACAGATTTTGGAAAGAGATCTTCACAAACAAACAGAC	684
QY	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPromet	217
Db	685	CTGGCTTCACACATGATAGACCAACACCCGCTCTGGGGCTGGCACCCCTGCTCCGATG	744
QY	218	CysLysGlySerArgCysTrrGlyGlnSerSerGluAspCysGlnSerLeuThrArgThr	237
Db	745	TGTAAAGGCTCCCTCCTCTGGGAGAGAGTTCGAGGATTTGTACAGCTCTGACGGGCACT	804
QY	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisAspIu	257
Db	805	GTCCTGTCCGGTGGCTGTGCCCGCTCGAAGGGGCCACTCCCACTATACGTCTGCATAG	864
QY	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277
Db	865	CAGTGTGTGCGGGCTCACGGGGCCCCAACACTCTGATGCTGGCCCTGCTCCACTTC	924
QY	278	AsnHisSerGlyIleCysGlyLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297
Db	925	AACCAACAATGGCACTGTGAGCTGCACTGCCACGCCCTGTCACTACAAACAAACAGAC	984
QY	298	PheGlnSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCysValThrAla	317
Db	985	TTTGTGTCATGACCCCAATCCCGAGGGCGGTATATACATTTGGCGGCGACTGTGACTGCC	1044
QY	318	CysProTyrAsnTyrIleAspSerThrAspValGlySer-----	329
Db	1045	TGTCCCTACACTACCTTTTCTACGGACGTGGATCTTGCAACCTGTGCCCCCTGCAC	1104
QY	329	-----	329
Db	1105	AACCAAGAGTGACAGACAGAGAGATGGAAACACACCGGTGTGAGAAAGTCAGTAAGCCCTGT	1164
QY	329	-----	329
Db	1165	GCCCCAGTGTCTATGTGTCTGGGCATGAGCACTTGCGAGAGGTGAGGGCACTTACCACT	1224
QY	329	-----	329
Db	1225	GCCAAATATCCAGAGATTTCGTGGCTGCAGAAATAATCTTTGGAGAGCCTGGCATTTCTGGCG	1284
QY	329	-----	329
Db	1285	GAGAGCTTTGATGGGAGCCAGGCTCCAAACACTGCCCGCTCCAGCCAGAGACAGCTCCAA	1344
QY	329	-----	329
Db	1345	GTTGTTGAGACTCTGGAAGAGATCACAGTTACTCTATACATCTCAGACATGGCCGGACAGC	1404
QY	329	-----	329
Db	1405	CTGCCTGACTCAGACGTCTTCCAGAAACCTGCAAGTATATCCGGGAGAGATTTCTGCACAAAT	1464
QY	329	-----	329
Db	1465	GAGCGCTACTGCTGACCTTCGCAAGGGCTGGGAGTCATGACTGGCTGGGGCTGCGCTCACTG	1524

QY 329 ----- 329
Db 1525 AGGAACTGGGAGTGAAGTGGCCCTCATCCACATTAACACCCACCTCTGCTGTCAC 1584
QY 329 ----- 329
Db 1585 ACGGTGCCCTGGGACAGCTCTTTCGGAACCCGACCAAGCTCTGCTCAACACTGCCAAC 1644
QY 329 ----- 329
Db 1645 CGGCAAGAGCAGAGTGTGTGGGGAGGGGCTGGCCCTGCCACAGCTGTGCGCCGAGGG 1704
QY 329 ----- 329
Db 1705 CACTGTGGGGTCCAGGGCCACCACAGTGTGTAAGTGCAGCAGTCTCTCGGGGCCAG 1764
QY 329 ----- 329
Db 1765 GAGTGGGTGAGGAATGCCAGTACTGCAAGGGGCTCCCGAGGAGTATGTGAATGCCAG 1824
QY 329 ----- 329
Db 1825 CACTGTTTGGCGTGCACCCCTGAGTGTACAGCCCGAGAATGGCTCAGTACCTGTTTGA 1884
QY 329 ----- 329
Db 1885 CCGGAGGCTGACCAAGTGTGTGGCCTGTGCCACTATTAAGACCCCTCTTCTGCGTGGCC 1944
QY 329 ----- 329
Db 1945 CGGTGCCCGACGGGTGTGAACCTGACCTCTCTACATGCCCATGTGGAAGTTTCCAGAT 2004
QY 329 ----- 329
Db 2005 GAGAGGGGCGCATGCGAGCTTGGCCCATCAAGTGCACCCACTCTGTGTGAACCTGGAT 2064
QY 329 ----- 329
Db 2065 GACAAGGGCTGCCCGCCGAGAGAGAGAGCCCTCTGACGTCCATCGTCTGCGGTG 2124
QY 329 ----- 329
Db 2125 GTTGGCATTTGCTGCTGTGTGTGTGTGGGGGTGCTTTTGGGATCCTCATCAAGCGACGG 2184
QY 329 ----- 329
Db 2185 CAGCAGAAATCCGAAAGTACAGATGCGGAGACTGCTGCAGAAACGAGAGTGTGGAG 2244
QY 329 ----- 329
Db 2245 CCGCTGACACCTAGCGGAGCGATGCCAACGAGCGCATGCGGATCTGAAAAGAGACG 2304
QY 329 ----- 329
Db 2305 GAGCTGAGAAAGTGAAGTGTGATCTGTGGGCTTTTGGCAGAGTCTCAAGGGGATC 2364
QY 329 ----- 329
Db 2365 TGGATCCCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGAGGAAAC 2424
QY 329 ----- 329
Db 2425 ACATCCCCCAAGCAAAAGAAATCTTAGACGAACATAGCTGATGCTGTGTGGGC 2484
QY 329 ----- 329
Db 2485 TCCCATATGTCTCCGCTTCTGGGATCTGCTGATCATCAGGATGACGCTGTGAC 2544
QY 329 ----- 329
Db 2545 CAGCTTATGCCCTATGTGCTCTTAGACATGTCCGGGAAACCGGAGCGCTGTGGG 2604
QY 329 ----- 329

Db 2605 TCCAGGACCTGTGAAGTGTGTATGACAGATTGCCAAGGGATGACCTACTGAGGAT 2664
QY 329 ----- 329
Db 2665 GTGGGCTCTTACACAGGACTTGGCCGCTCGAAGCTGCTGGTCAAGAGTCCCAACAT 2724
QY 329 ----- 329
Db 2725 GTCAAAATTAAGACTTCGGGCTGGCTCGGCTGTGACATTGACAGACAGATACAT 2784
QY 329 ----- 329
Db 2785 GCAGATGGGGCAAGGTGCCCATCAAGTGAATGCGCTGAGTCCATTCTCCGCGCGG 2844
QY 329 ----- 329
Db 2845 TTCACCCACAGAGTGTGTGTGAGTTATGTGTGACTGTGTGGAGTGAATCTTTT 2904
QY 329 ----- 329
Db 2905 GGGCCAAACCTTACGATGGATCCAGCCGGGAGATCCCTGACTGTGAAAAGGG 2964
QY 329 ----- 329
Db 2965 GAGCGGCTGCCAGCCCGCATTCGACATTGATGTATGATGATCATGTCAAAATGT 3024
QY 329 ----- 329
Db 3025 TGGATGATTACTCTGAATGTGGCCCAAGATCCGGAGTGTGTGTGAATTTCCCGC 3084
QY 329 ----- 329
Db 3085 ATGGCCAGGAGCCCGCGCTTTGTGTATCCAGAAATGAGACTTGGCCAGCAGT 3144
QY 329 ----- 329
Db 3145 CCCTTGACAGACCTTCTACCCCTCAGCTGTGAGAGCATGATGGGGACCTGGTG 3204
QY 330 ----- Gly 330
Db 3205 GATGCTGAGAGTATCTGTGATCCCGACAGGGCTTCTTGTCCAGACCTGCCCGCGGC 3264
QY 331 AlAGlyGlymeValHisHisArgSerSerThrArgSerGlyGlyLeu 330
Db 3265 GCTGGGGCATGTGCTCCACAGAGCCGACAGCTCATCTACAGAGTGGCGGGGAGC 3324
QY 331 LeuThrLeuGlyLeuGlyProSerGlyGlyAlaLeuProArgSerProLeuAlaProSer 370
Db 3325 CTGACCTAGAGGCTGGAGCCCTGTGAAGAGAGAGCCCGCAGGTCTCCACTGGCACCTCC 3384
QY 331 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyAlaAlaLeuGlyLeu 390
Db 3385 GAAGGGCTGGCTGCCATGATTTGATGTGACTGGGAATGGGGGACCAAGGGGCTG 3444
QY 391 GluSerLeuProThrHisAspProSerProLeuGlyAlaArgTySerGlyAspProThrVal 410
Db 3445 CAAAGCCTCCCAACATGAGCCCGACCTCTACAGGGGACAGTGTGGAGCCCGCACGTA 3504
QY 411 ProLeuProSerGlyThrAspGlyTyValAlaProLeuThrCysSerProGluProGlu 430
Db 3505 CCCCTGCCCTGTGAGACTGATGGCTTACGTTGCCCTTGACCTGCAGCCCCCAGCTGAA 3564
QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuPro 450
Db 3565 TATGTGAACAGCAGATGTGGGCCACCGCCCTTCCGCGAGAGGGGCGCTGCTCCT 3624
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGlyAlaArgTySerProGlyGlyAsn 470
Db 3625 GCTGCCCGACCTGTGTGCTGTGAAAGGGCCAAAGACTCTTCCCGAGGGAAGAAAT 3684
QY 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490

Db 3685 GGGGCTGTCMAAGACGTTTTCCTTTGGGGGCTGCCGTGGAGAACCCCGAGTACTTGACA 3744
 Qy 491 ProGlnGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheasp 510
 Db 3745 CCCCAGGAGAGAGCTGCCCTTACGCCCACTCTCTGCTTACGCCCACTCTGAC 3804
 Qy 511 AsnLeuTYTTPRAspGlnAspProProGlnArgGlyAlaProPheSerThrhPheLys 530
 Db 3805 AACCTCTATTACTGGACAGACACACAGAGGGGGGCTCCACCCAGCACTTCAAA 3864
 Qy 531 GlyThrProThrAlaGlnAsnProGluTYTLeuGlyLeuAspValProAlaAlaPro 550
 Db 3865 GGGACACCTACGCGAGAGAACCCAGATACGCGCTGAGAGTCCCA-----CTGTGA 3918
 Qy 551 AlaArgSerProSerPro 556
 Db 3919 ACCAGAGGCCCAAGTCCG 3936
 RESULT 11
 ABK14057
 ID ABK14057 standard; DNA; 9274 BP.
 AC ABK14057;
 DX 23-APR-2002 (first entry)
 XX
 DE Human HER2 (ErbB2) transgene plasmid construct DNA.
 XX
 KW Human; HER2; ErbB; epidermal growth factor receptor; gene; ds;
 KW anti-ErbB antibody-maytansinoid conjugate; cancer; tumour; breast; ovary;
 KW stomach; endometrium; salivary gland; lung; kidney; colon; colorectum;
 KW thyroid; pancreas; prostate; bladder; ErbB2; neuronal disorder;
 KW glioma disorder; astrocytoma disorder; hypothalamic disorder;
 KW glandular disorder; macrophage disorder; epithelial disorder;
 KW stromal disorder; blastocoele disorder; inflammatory disorder;
 KW angiosarcoma disorder; immunological disorder.
 KW
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN US200201587-A1.
 XX
 PD 03-JAN-2002.
 PF 16-MAR-2001; 2001US-0811123.
 PR 16-MAR-2000; 2000US-189844P.
 PR 05-OCT-2000; 2000US-238327P.
 XX
 PA (ERIC/) ERICKSON S.
 PA (SCHW/) SCHWALL R.
 PA (SLIW/) SLIWOMSKI M.
 XX
 PI Erickson S, Schwall R, Sliwowski M;
 DR WPI: 2002-163686/21.
 DX
 PT Treating tumour characterised by overexpression of epidermal growth
 PT factor receptor, ErbB or cancer in mammal, comprises administering
 PT anti-ErbB antibody-maytansinoid conjugate to the mammal
 XX
 PS Example 3; Fig 7; 93pp; English.
 CC The invention relates to treating a tumour in a mammal, where the tumour
 CC is characterised by the overexpression of an epidermal growth factor
 CC receptor (ErbB) and does not respond or responds poorly, to treatment
 CC with an anti-ErbB antibody, comprising administering to the mammal an
 CC anti-ErbB antibody-maytansinoid conjugate. The method is useful for
 CC treating cancer or tumours of the breast, ovary, stomach, endometrium,
 CC salivary gland, lung, kidney, colon, colorectum, thyroid, pancreas,
 CC prostate and bladder, preferably breast cancer. The breast cancer is a
 CC metastatic breast cancer or an aggressive form of metastatic breast
 CC cancer which overexpresses ErbB2. The method is also useful for treating

CC neuronal, glial, astrocytic, hypothalamic, glandular, macrophage,
 CC epithelial, stromal, blastocoele, inflammatory, angiosarcoma
 CC immunological disorders. This sequence represents a HER2 (ErbB2)
 CC transgene plasmid construct used to direct expression of native human
 CC HER2 in the mammary gland of a transgenic mouse.
 XX
 SQ Sequence 9274 BP; 2152 A; 2538 C; 2489 G; 2095 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.76e-97 Length: 9274
 Score: 2422.00 Matches: 528
 Percent Similarity: 41.12% Conservative: 9
 Best Local Similarity: 40.43% Mismatches: 31
 Query Match: 64.78% Indels: 738
 DB: Gaps: 6
 US-09-821-883-2 (1-690) x ABK14057 (1-9274)
 Qy 2 ArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 21
 Db 1695 CGCCCGCGGCGCTCCAGCGGGGTCCAGCGGAGCCATGAGACTGGCG-----GCC 1745
 Qy 22 LeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSer 41
 Db 1746 TTGTGCGCCCTGG-----GGGCTCCTCTCCGCGCTCTTCCCGCCGAGCCCGAGC 1796
 Qy 42 ThrGlnValCysThrGlyTYTThrAspMetLysLeuArgLeuProAlaSerProGluThrHis 61
 Db 1797 ACCCAAGTGTGACCGCGCCACACATGAAAGTGGCGCTCCGCGAGTCCAGACCCCGC 1856
 Qy 62 LeuAspMetLeuArgHisLeuTYTLeuGlnGlyCysGlnValValGlnGlyAsnLeuGluLeu 81
 Db 1857 CTGGACATGCTCCGCGACCTTACACAGGGCTCCAGGGTGTGTCCAGGAACTGGAGACTC 1916
 Qy 82 ThrTYTLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGlnGly 101
 Db 1917 ACCTACCTGCGCCACCAATGCCAGCTGCTTCCTGTCAGATATCCAGAGGTGACAGGC 1976
 Qy 102 TYTValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgGluArgIleVal 121
 Db 1977 TACGTGCTCATGCTGCACACCAAGTAGAGCGTCCCACTCCAGAGCTGGGATTGG 2036
 Qy 122 ArgGlyThrGlnLeuPheGlnAspAsnTYTAlaLeuAlaValLeuAspAsnGlyAspPro 141
 Db 2037 CGAGGACCCACAGCTCTTGAGAGACAATGATGCCGTGGCGCTGTAGACAATGAGACCGC 2096
 Qy 142 LeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuLeu 161
 Db 2097 CTGAACAAATACCAACCCCTGTCCAGAGGGCTCCCGAGAGGCTCGGAGCTGACGCT 2156
 Qy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAspProGlnLeuLys 181
 Db 2157 CGAAGCTTCCAGAGATCTTGAAGAGAGGGGCTTGTATCCAGAGGACCCCAAGCTTGC 2216
 Qy 182 TYTAlaSerThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
 Db 2217 TACAGAGACAGATTTTGGAGAGACATCTTCCAAAGAAACAACAGCTGGCTGCACA 2276
 Qy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
 Db 2277 CTGATATAGACACCAACCGCTGCGGCTGCGCCCTGTTCTCCGATGTGTAAAGGCTTC 2336
 Qy 222 ArgCysTrpGlyGlySerSerGlnAspCysGlnSerLeuThrArgThrValLysAlaGly 241
 Db 2337 CCCTGCTGGGAGAGAGATTCTTGAAGATTGTCAAGCTTGAACCGCACTGTCTGTGCGGT 2396
 Qy 242 GlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAlaAla 261
 Db 2397 GGCTGTGCGCGGTGCAAGGGGCACTGCCCACTGATGTGTCAGAGAGAGTGTGTCG 2456
 Qy 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281
 Db 2457 GGCTGACAGGGGCCCAAGACATCTGACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 2516

Qy 282 IlleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMet 301
Db 2517 ATCTGTAGCTGCACCTGCCAGCCCTGTGCTACACACAGACAGCTTTGATGTCATG 2576
Qy 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321
Db 2577 CCCAATCCCGAGGGCCGGTATACATTCGGCCAGCTGTGTACTGCTGCTCCCTACAC 2636
Qy 322 TyrlleuSerThrAspValGlySer 329
Db 2637 TACCTTTCTACGACGTGGATCTTCGACCCCTGCTGCCCCCTGCACACCAAGAGGTG 2686
Qy 329 329
Db 2697 ACACAGAGAGTGAACACACAGCGGTGTGAGAACTGCAGCAAGCCCTGTGCCCGAGTGTGC 2756
Qy 329 329
Db 2757 TATGCTGTGGGAGAGAGACACTTGCAGAGGTGAGGCGAGTTACAGTGCCTATATCCAG 2816
Qy 329 329
Db 2817 GAGTTTCTGGCTGCAAGAAAGATCTTTGGAGCCTGCGCATTTCTCCGGAGAGCTTTGAT 2876
Qy 329 329
Db 2877 GGGAGCCAGCCTCCACACTGCCCCGCTCCAGCCAGACAGCTCAAGTGTGAGACT 2936
Qy 329 329
Db 2937 CTGGAAGAGATCACAGGTACCTATACATCTCAGCATGCGCGGAGACGCTGCTGACCTC 2996
Qy 329 329
Db 2997 AGCCTCTCCAGAACTTCAGATATCCGGGAGCAATTCTGCACATGGCGCTACTCG 3056
Qy 329 329
Db 3057 CTGACCTGCAAGGCTGGGCTATCAGCTGGCTGGCGCTCAGTGAAGGAACTGGGC 3116
Qy 329 329
Db 3117 AGTGGACTGGCCCTCATTCACACCATACACCCACTGCTTCTGTCACAGGTGCCCTGG 3176
Qy 329 329
Db 3177 GACACGCTTTTGGAAACCGGACCAAGCTGTCTGCACACTGCCAACCGGCCAGAGGAC 3236
Qy 329 329
Db 3237 GAGTGTGGGCGAGGGCTGGCTGCCACAGCTGTGCCCGGAGGCACTGCTGGGGT 3296
Qy 329 329
Db 3297 CCAAGGCCCAACCACTGTCTCACTGCAGCCAGTTCTTGGGGGCCAGAGTGGTGAG 3356
Qy 329 329
Db 3357 GAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGGCACTGTTGCCG 3416
Qy 329 329
Db 3417 TGCCACCTGAGTGTCAAGCCCAAGAAATGCTCAGTGAAGCTTTTGGACGGAGGCTGAC 3476
Qy 329 329
Db 3477 CAGTGTGGCTGTGCCCACTATAGGACCTTCTTGTGCTGCGCCGCTGCCAGC 3536
Qy 329 329
Db 3537 GGTGTGAACCTGACCTCTCTCATATGCCCATCTGGAAGTTTCCAGATGAGAGGGGCA 3596

Qy 329 329
Db 3597 TGCCAGCCTTGCCCATCACTGCACCCACTCTGTGTGACCTGATGACAAAGGCTGC 3656
Qy 329 329
Db 3657 CCCCGCAGAGAGAGAGCCAGCCCTGTGACGTCCATCTCTGCGGTGGTTGGCATTCG 3716
Qy 329 329
Db 3717 CTGTGTGTGCTTGGGGTGTGCTTTGGATCTCATCAAGCGAGCGCAAGAAATC 3776
Qy 329 329
Db 3777 CGAAGTACAGATGCGGAGACTGCTGCAGAAAGGAGCTGTGAGCCGCTGACACT 3836
Qy 329 329
Db 3837 AGCGAGCCGATGCCCAACAGGCGCAGATGCGATCTGAAAGAGAGAGCTGAGGAG 3896
Qy 329 329
Db 3897 CTGAAGTGTGTGATCTGGCGCTTTTGGCAGCTACAAAGGCACTGTGATCCCTGAT 3956
Qy 329 329
Db 3957 GGGAGATGTGAATAATTCAGTGGCCATCAAAAGTGTGAGGAAACACATCCCCAAA 4016
Qy 329 329
Db 4017 GCCAACAAGAAATCTTAGACGAACATACGTATGCTGTGTGGCTCCCATATGTC 4076
Qy 329 329
Db 4077 TCOCGCTTGTGGGCTATCTGCTGACATCCACGCTGAGCTGTGACACAGCTTATGCC 4136
Qy 329 329
Db 4137 TATGGCTGCTTTAGACATGTCCGGGAAACCGCGAGCCTGGCTCCAGACCTG 4196
Qy 329 329
Db 4197 CTGAAGTGTGTATGACGATTTGCCAAGGAGTGAAGTACCTGGAGATGTGCGCTGTA 4256
Qy 329 329
Db 4257 CACAGGACTTGGCGCTCGGAACGTGCTGTCAAGAGTCCCAACATGTCAAAATTACA 4316
Qy 329 329
Db 4317 GACTTCGGGCTGGCTCGGCTGTGACATTTGACGAGACAGAGTACATGAGATGGGGC 4376
Qy 329 329
Db 4377 AAGTGCCTCAAGTGAATGAGTGGCTGAGTCCATTCTCCGCGCGGCTTCCACCACAG 4436
Qy 329 329
Db 4437 AGTATGTGAGAGTATATGTGATCTGTGAGAGCTGATGACTTTTGGGCCCAACT 4496
Qy 329 329
Db 4497 TACATGGATGCCAGCGCGGAGATCCGTACCTGTGAAAGGGGAGCGGCTGCC 4556
Qy 329 329
Db 4557 CAGCCCCCATCTGCACCATTTGATGTCTACATGATCATGTGCAAAATGTGATGATTGAC 4616
Qy 329 329
Db 4617 TCTGATGTGGGCAAGATTCGGGAGTGTGTCTGAATTCGCCCATGGCCAGGAG 4676
Qy 329 329


```

Db 4677 CCCAGCCTTTGGTGCATCCAGATGAGACTTGGGCCAGCAGCTCCCTGGACAGC 4736
OY 329 -----
Db 4737 ACCTTCAACCGCTACTGCTGAGAGACATGATGGGGACCTGGTGAATGCTGAGAG 4796
OY 330 -----GlyAlaGlyGlyMet 334
Db 4797 TATCGTATACCCAGCAGAGGCTTCTTCTGTCACAGACCTGCCCGGGGCGCTGGGCGATG 4856
OY 335 ValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeuThrLeuGly 354
Db 4857 GTCCACACAGAGCAGCAGCAGCTCATCTACAGAGAGTGGCGGGGAGCCTGACACTAGAG 4916
OY 355 LeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGluGlyAlaGly 374
Db 4917 CTGGAGCCCTCTGAAGAGAGAGGCCCCAGGTCTCCACTGGGACCTCCGAAGGGGCTGGC 4976
OY 375 SerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeuGlnSerLeuPro 394
Db 4977 TCCGATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5036
OY 395 ThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValProLeuProSer 414
Db 5037 ACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGAGACCCACACAGTACCCCTGCCCT 5096
OY 415 GluThrArgGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAlaGln 434
Db 5097 GAGACTGATGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5156
OY 435 ProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAlaArgPro 454
Db 5157 CCAGATGCTTGGGCCCCAGCCCTTGGCCCGAGAGGGGCCCTGCTGCTGCTGCTGCTGCTGCT 5216
OY 455 AlaGlyAlaThrLeuGluArgAlaAlaGlyThrLeuSerProGlyLysAsnGlyValValLys 474
Db 5217 GCTGGTGCACCTCTGAAAGGGGCCAAGACTCTCTCCACAGGAGAAATGGGGTCTGTCANA 5276
OY 475 AspValPheAlaPheGlyGlyAlaValAlaGluAsnProGluTyrLeuThrProGlnGly 494
Db 5277 GACGTTTTTGGCTTTGGGGGGTCCCTGGAGAAACCCCGAGTACTTACACCCAGGAGAGGA 5336
OY 495 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyrTyr 514
Db 5337 GCTGCCCTCAGCCCCCAGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5396
OY 515 TrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThrProThr 534
Db 5397 TGGGACACAGAGACCCAGAGGGGGGCTCCACCCAGCAGCAGCTTCAAGGGAGACACTAGC 5456
OY 535 AlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProAlaArgSerPro 554
Db 5457 GCAGAGAACCCAGACTGCTGGTCTGAGCGTCCA-----GTGTGAACCCGAAGAGGCA 5510
OY 555 SerProSerThrGlnProTropGluHisValAsnAlaIleGlnGluAlaAlaArgArgLeu 574
Db 5511 AGTCGGCAG-----AAGCCTGATGTCTCTCAGG-----GAGCAGAGGAAGGGGCTCTG 5561
OY 575 AsnLeuSerArgAspThr 580
Db 5562 AGCTATTCAGAACTAGT 5579

```

RESULT 12

ID AA14812 standard; cDNA: 4472 BP.

XX AC AA14812;

XX DT 08-AUG-2000 (first entry)

XX DE cDNA encoding the SPLICE erB-2 receptor protein.

```

KW SPLICE erB-2 receptor protein; cell transformation disorder; cancer;
KW tumor cell proliferation; tissue degeneration; arthropathy;
KW bone resorption; inflammatory disease; degenerative disorder;
KW wound healing; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 175..3942
XX FT /*tag=a
XX FT /product= "SPLICE erB-2 protein"
XX MO200020579-A1.
XX PD 13-APR-2000.
XX PF 01-OCT-1999; 99WO-CA00912.
XX PR 02-OCT-1998; 98US-0165192.
XX PA (UVMC-) UNIV MCMASTER.
XX PI Muller WJ, Siegel PM;
XX WP1; 2000-303768/26.
XX DR P-PSDB; AAY84780.
XX PT Nucleic acid encoding an erB-2 receptor protein designated SPLICE
XX erB-2, inhibitors of the protein are useful for treatment of cancer -
XX Claim 4: Fig 1: 60pp; English.
XX PS
XX CC The present sequence encodes a SPLICE erB-2 receptor protein. The
XX protein has an in-frame deletion of 16 amino acids, 2 of which are
XX conserved cysteine residues, compared to the unspliced protein. The
XX erB-2 polynucleotide is used to construct probes for detecting
XX disorders of cell transformation such as cancer. Antibodies to the
XX protein may be used to detect SPLICE erB-2 in a sample. Agents
XX (e.g. antisense oligonucleotides) which inhibit the expression of
XX SPLICE erB-2 are useful for reducing tumor cell proliferation and
XX treating cancer. Substances which stimulate SPLICE erB-2 are useful
XX for treating conditions involving damaged cells including conditions
XX in which degeneration of tissue occurs, such as arthropathy, bone
XX resorption, inflammatory diseases, degenerative disorders of the
XX central nervous system and wound healing.
XX SQ Sequence 4472 BP; 902 A; 1383 C; 1328 G; 859 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.33e-97 Length: 4472
XX Score: 2418.00 Matches: 523
XX Percent Similarity: 41.06% Conservative: 22
XX Best Local Similarity: 40.67% Mismatches: 736
XX Query Match: 64.67% Indels: 736
XX DB: 21 Gaps: 4
XX
XX US-09-821-883-2 (1-690) x AA14812 (1-4472)
XX OY 2 ArgAlaAlaProLeuLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 17
XX Db 118 CGCGCCCTCCAGCGCGGTCCAGCGGAGCCATGGGGCCGAGCGCAGTGAACCATG 177
XX OY 18 PheLeuPheLeuLeuPhePheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
XX Db 178 GAGCTGGCGGCTTGTGCCGCTGG-----GGGCTCTCTCTGCGCTTGGCCCCC 228
XX OY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
XX Db 229 GGAGCGCGGAGAGCAGCAAGTGTGACACGCGACAGATGAAGACTGCGGTCTCCCTCAGT 288
XX OY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
XX Db 289 CCGGAGACCCACCTGAGATGCTCGCGCACCTCTACAGAGGGGCTGCCAGGTGTGACAGGA 348

```

Qy	78	AsnLeuGluLeuThrTyrIleuProThrAsnAlaSerLeuSerPheLeuGlnAsnIleGln	97
Db	349	AACCTGGAACTCACCTTACCTGCGCCACCAATGCCAGCCTGTCTTCCGACGAGATATCCAG	408
Qy	98	GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
Db	409	GAGGTGCAGGGCTACCTGCTCATCGCTCACACCAAGTGAAGGAGGCTCCACATGCAGAGG	468
Qy	118	LeuArgIleValAlaArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAsp	137
Db	469	CTGGCGATTGTGGAGGCAACCCAGCTCTTGGAGACAACTATGCCCTGGCCGTGCTAGAC	528
Qy	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
Db	529	AATGAGACCCCGCTGAACAATACACCCCTGCTCACAGGGGCTCCCGAGGAGGCTGCGG	588
Qy	158	GluLeuGlnLeuArgSerIleuThrGluIleuGlyGlyValLeuIleGlnArgAsn	177
Db	589	GAGCTGCAGCTTCGAAGCCCTCACAGAGATCTTGAAGAGAGGGCTTGTATCCAGCGAAC	648
Qy	178	ProGlnLeuCysTyrGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGln	197
Db	649	CCCGACCTGTGCTACACAGACAGATTTGTGGAAGAGACTCTTCCACAGAACACACAG	708
Qy	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPromet	217
Db	709	CTGGCTCTACACTGTATTAACACCAACCGCTCTCGGGCCCTCCACCCCTTCTCCGATG	768
Qy	218	CysLysGlySerArgTyrPrgLysIleuSerSerGlnAspCysGlnSerLeuThrArgThr	237
Db	769	TGTAAAGGGTCCCGCTCGCTGGGGAAGAGTCTGAGAGATTGTACAGCCGACGCGCACT	828
Qy	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu	257
Db	829	GCTGTGCGCGGTGCTGCTCCGCTCCAAAGGGGCCACTGCCACTACTGCTCCCATAG	888
Qy	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaLysPhe	277
Db	889	CAGTGTGCTCGCGGTGCTGCAAGGGGCCCAAGCACTGTGACTGCTGGCTGCTCCACTTC	948
Qy	278	AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297
Db	949	AACCCAGATGGCATCTGTAGTGCACCTGCCAGCCCTGGTCACTAACACACAGACAG	1008
Qy	298	PheGluSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCysValThrAla	317
Db	1009	TTTGAATCCATGCGCAATCCCGAGGGCGGTATATCATTCGGCGCCAGCTGTGTACTGCC	1068
Qy	318	CysProTyrAsnTyrIleuSerThrAspValGlySer	329
Db	1069	TGTCCCTACAGTACTCTTCTAGCGAGCGTGGATCTCTGCACCCCTGTGCCCCCTGCAC	1128
Qy	329	-----	329
Db	1129	AACCAAGAGTGACACAGAGATGGAGACACACAGCGGTGTGAGAAATGCAAGACCCCTGT	1188
Qy	329	-----	329
Db	1189	GCCCGAGTGTGCTATGTGCTGGGCAATGAGCACTTGCAGAGGTGAGGGCATTACCACT	1248
Qy	329	-----	329
Db	1249	GCCAAATATCCAGAGATTGTCTGGCTGCAAGAAGATCTTTGGGAGCCTGGCATTTCTGCCG	1308
Qy	329	-----	329
Db	1309	GAGAGCTTTGATGGGAGCCAGCCTCCAACTGCCCCCTCCAGCCAGAGAGACTCAA	1368
Qy	329	-----	329
Db	1369	GTTGTTGAGACTCTCGAAGAGATCACAGGTTTACCTTATACATCTCACAGCATGGCCGGACAG	1428

```

Db 2509 TCCCATATGTCCTCCGCTTGTGGGCAATGCTGACATCCAGCGTGGAGCTGTGACA 2568
Oy 329 ----- 329
Db 2569 CAGCTTATGCCCTATGCGCTTGTAGACATGTCGGGAAACCGGGGACGCTGGGC 2628
Oy 329 ----- 329
Db 2629 TCCGAGACCTGCTGACTGGTGTATGCAAGATTGCCAAGGGGATGAGCTACTGGAGAT 2688
Oy 329 ----- 329
Db 2689 GTGCGCTCGTACACAGGACTTGGCCGCTCGAAGCTGCTGTCAGAGATCCAAACAT 2748
Oy 329 ----- 329
Db 2749 GTCAAAATTACAGACTTGGGGCTGGCTGGCTGAGCATTTGACAGACAGATACAT 2808
Oy 329 ----- 329
Db 2809 GCAGATGGGGGCAAGTGGCCCATCAAGTGGATGGCGCTGGAGTCCATTCTCCGGCGG 2868
Oy 329 ----- 329
Db 2869 TTCACCAACAGATGATGTGTGAGATTATGTTGACTGTGTGGAGCTGATGACTTTT 2928
Oy 329 ----- 329
Db 2929 GGGGCCAAACCTTACAGATGGATGCCAGCCGGAGATCCCTGACTGCTGGAAAAAGGG 2988
Oy 329 ----- 329
Db 2989 GAGCGCTGCCCGACCCCGCATCTGCACATTGATGCTACATGATCATGCTCAAAATG 3048
Oy 329 ----- 329
Db 3049 TGGATGATTGACTGTGAATGTGCGCCAAAGATCCGGAGTTGGTGTGAATTTCTCCGC 3108
Oy 329 ----- 329
Db 3109 ATGCCAGGAGACCCCGACGGCTTTGTGTGTCATCCAGAAATGAGACTTGGGCCACCACT 3168
Oy 329 ----- 329
Db 3169 CCCTTGAGACGACCTTCTACCGCTCAGCTGCTGAGAGACATGAGGGGACCTGCTG 3228
Oy 330 -----Gly 330
Db 3229 GATGCTGAGAGATATCTGTGACCCAGAGGGCTTCTGTGCAAGCCCTGCCCGGGC 3288
Oy 331 AAGAGlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyLysP 350
Db 3289 GCTGGGGGATGCTCCACACAGGACCGCACCTCATCTACAGAGATGGCGGTGGGGC 3348
Oy 3351 LeuThrLeuGlyLeuGlyLysProSerGluGluGluGlyMetGlyAlaAlaLysGlyLeu 370
Db 3349 CTGACACTAGAGGCTGAGGCCCTCTCTAAGAGAGAGGCCCGGCTCTCCACTGGCACCTTC 3408
Oy 371 GlnGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 390
Db 3409 GAAAGGGGCTGGCTCGATGATTTGATGTTGATGCTGAGGAAATGGGGGCAAGGAGGCTG 3468
Oy 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410
Db 3469 CAAGACCTTCCTCCACATGACCCCGCCTCTACAGCGGTACAGTGAAGACCCACAGTA 3528
Oy 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu 430
Db 3529 CCCCTGCGCTCTGACAGCTATGCTACGTTGCGCCCTGACTGCAAGCCCAAGCTGAA 3588
Oy 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyLysProLeuPro 450

```

```

Db 3589 TATGTGAACACAGCATGTTGGCCCGCCAGCCCTTCGCCCGAGAGAGGCGCTCTGCT 3648
Oy 451 AlaAlaArgProAlaGlyAlaLeuThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn 470
Db 3649 GCTGCCCGACCTGCTGTGCTGCTGCAAGAGCCCAAGACTCTCTCCACAGGAAGAT 3708
Oy 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValAlaGluAsnProGlyLysThr 490
Db 3709 GGGGTGCTCAAGAGACTTTTCCCTTGGGGGTGCGCGTGAGAACCCGAGACTTTCACA 3768
Oy 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
Db 3769 CCCGAGGAGAGCTGCCCTCAGCCCACTCTCTGCTGCTTCCAGCCACAGCTTTCAGC 3828
Oy 511 AsnLeuTyrTyrThrAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
Db 3829 AACCTCTATTACTGGACAGGACACACAGCGGGGGCTCCACCGACCTTCAAA 3888
Oy 531 GlyThrProThrAlaGluAsnProGlnLysLeuAspValProAlaAlaPro 550
Db 3889 GGGACACCTAGCGGCGAGACACCCAGATACCTGGGTCTGGAGCTGCCA-----GTGCA 3942
Oy 551 AlaArgSerProSerPro 556
Db 3943 ACCAGAGGCCCAAGTCCC 3960

RESULT 13
AAZ31071
ID AAZ31071 standard; DNA; 4473 BP.
AC AAZ31071;
XX 17-JAN-2000 (first entry)
DT 17-JAN-2000 (first entry)
DE HER-2 nucleic acid sequence.
XX HER-2: c-neu; ErbB2; transmembrane receptor; tyrosine kinase activity;
XX epidermal growth factor receptor; EGFR; HER-1; cancer; breast cancer;
XX ovarian cancer; gastric cancer; antisense oligonucleotide; expression;
XX hyperproliferative disease; ss.
XX Homo sapiens.
XX OS
XX MO3948906-A1.
XX 25-MAR-1999; 99WO-US06492.
XX 26-MAR-1998; 98US-0048804.
XX (ISIS-) ISIS PHARM INC.
XX (PENN-) PENN STATE RES FOUND.
XX Bennett CF, Lipton A, Wilters LM;
XX WPI: 1999-610749/52.
XX New antisense sequences used to treat hyperproliferative conditions,
XX especially cancer -
XX Examples: Page 38-39; 44pp; English.
XX
XX This is the human HER-2 polynucleotide sequence. The HER-2 gene also
XX called c-neu and ErbB2, encodes a transmembrane receptor, with tyrosine
XX kinase activity. HER-2 is related to the epidermal growth factor
XX receptor (EGFR or HER-1). Aberrant HER-2 expression is present in a wide
XX number of cancers, especially breast, ovarian and gastric cancers. This
XX sequence is used in the invention to design 12-25 nucleotide
XX oligonucleotides that decrease the expression of human HER-2. The
XX oligonucleotides of the invention (AAZ31067-231070) can also be used for
XX modulating the expression of human epidermal growth factor receptor. The
XX oligonucleotides are used to treat diseases or conditions associated with

```



```
QY 329 ----- 329
Db 1609 ACGGTGCCCTGGAGCCACACTCTTTGGAAACCCGCCAACAGCTGCTGCTCCACACTGCCAAC 1668
QY 329 ----- 329
Db 1669 CGGCCAGAGCAGATGTGTGGCGAGGGCTGGCTGCCACAGACTGTGCCCGAGAGG 1728
QY 329 ----- 329
Db 1729 CACTGCTGGGGTCCAGGGCCCAACCAAGTGTCAACTGCACCCAGCTCTCGGGGCGAG 1788
QY 329 ----- 329
Db 1789 GAGTGGCGTGAAGATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAG 1848
QY 329 ----- 329
Db 1849 CACTGTTTGGCGTCCACCTGAGTGTGAGCCCGAAGATGGCTCAGTGACCTGTTTGA 1908
QY 329 ----- 329
Db 1909 CCGGAGCTGAACAGTGTGGCTGTGCCACTATAAGAACCTCCCTTGTGCGTGGCC 1968
QY 329 ----- 329
Db 1969 CGGTGCCCCAGGGGTGAAACGTGACCTTCCTACATGCCCATGTGAAGTTTCCAGAT 2028
QY 329 ----- 329
Db 2029 GAGGAGGCGCATGCCACCTTGCCCATCAACTGCACCCACTCCTGTGTGAGACTGGAT 2088
QY 329 ----- 329
Db 2089 GACAAAGGCTGCCCCCGGAGCAGAGAGCAGCCCTTGAGTCCATCATCTCTGCGGTG 2148
QY 329 ----- 329
Db 2149 GTTGGCATTCTGCTGTGTGTGTGTGGGGTGTCTTTGGGATCCTCATCAAGCGAGCG 2208
QY 329 ----- 329
Db 2209 CAGCAGAAATCCGGAGTACAGATGCGGAGACTGTCTCAGAAACGAGGTGTGGAG 2268
QY 329 ----- 329
Db 2269 CCGGTGACACCTAGCGGAGCGATGCCCAACAGCGCAGATGCCATCCTGAAGAGACG 2328
QY 329 ----- 329
Db 2329 GAGCTGAGGAAGTGAAGTGTGGATCTGGCGCTTTTGGCACAGTCTACAAGGCGATC 2388
QY 329 ----- 329
Db 2389 TGGATCCCTGATGGGAGAAATGTGAAATTCAGTGGCCATCAAGTGTGAGGGAAC 2448
QY 329 ----- 329
Db 2449 ACATCCCCCAAGCCAACAAAGAAATCTTAGACGAAGCATAGCTGATGGCTGTGGGC 2508
QY 329 ----- 329
Db 2509 TCCCATATGTCTCCCGCTTGTGGCATGTGCTGACATCCAGGGTGCAGCTGCTGACA 2568
QY 329 ----- 329
Db 2569 CAGCTTATGCCCTATGCTGCTTGAACCATGTCCGGAACCCGCGAGCGCTTGGGC 2628
QY 329 ----- 329
Db 2629 TCCAGAGACTGCTGAAGTGTGTATGAGATTGCCAAGGGGAGTACTACTGGAGAT 2688

QY 329 ----- 329
Db 2689 GTGGCGCTGTTACACAGGAGACTTGGCCCTGGAAACGTGTGTCAAGTGTCCAAACAT 2748
QY 329 ----- 329
Db 2749 GTCAAAATTACAGACTTTCGGCTGCGTCCGCTGTGACATTGACAGACAGATACCAT 2808
QY 329 ----- 329
Db 2809 GCAGTGGGGGCAAGTGGCCCATCAAGTGAATGGCGCTGAGTCCATTCTCCCGCGCGG 2868
QY 329 ----- 329
Db 2869 TTCAACCCACAGAGTGAATGTGGAGTATGTGTGACTGTGTGGAGCTGATCTTT 2928
QY 329 ----- 329
Db 2929 GGGGCCAAACCTTACATGATGGATCCAGCCCGGAGATCCCTGACCTTGGAAGGGG 2988
QY 329 ----- 329
Db 2989 GAGCGGCTGCCCAAGCCCCCATCTGCACATTGATGTCTACATGATCATGTCAATGT 3048
QY 329 ----- 329
Db 3049 TGGATGATTGACTGTGAATGTGCGGCAAGATTCCGGAGTTGTGTGTAATTCCTCCGC 3108
QY 329 ----- 329
Db 3109 ATGGCCAGGAGCCCCCAGCGCTTGTGTGTCATCCAAATGAGACTTGGGCCAGCCAGT 3168
QY 329 ----- 329
Db 3169 CCCTTGGACAGACCTTACCCTCTACCTGCTGAGAGACGATGACATGGGGACTGGTG 3228
QY 330 ----- 330
Db 3229 GATGCTGAGAGTATCTGTGTAACCCAGCAGGCGCTTCTGTCCAGACCTTGCCTCCGCGG 3288
QY 331 ----- 331
Db 331 ALaGlyMetValHisIleArgHisArgSerSerThrArgSerGlyIleIys 3350
Db 3289 GCTGGGGCATGTGTCACACAGGACCGCAGCTCATCTTACAGAGTGGCGGGGAGC 3348
QY 331 ----- 331
Db 331 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 3370
Db 3349 CTGACACTAGGGCTGAGACCTCTGAAGAGAGGCCCCCAGGTCTCCACTGGACCTTCC 3408
QY 371 ----- 371
Db 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 3390
Db 3409 GAAGGGCTGGTCCGATGTAATTTGATGTGACTGGGAATGGGGGAGCCAAAGGGGCTG 3468
QY 391 ----- 391
Db 391 GlnSerLeuProThrHisAspProSerProLeuGlnArgIleGlySerGluAspProThrVal 410
Db 3469 CAAGCCCTCCCCACACATGACCCCGCCCTTACAGCGGTACAGTGAAGAGACCCCACTA 3528
QY 411 ----- 411
Db 411 ProLeuProSerGluThrAspGlyThrValAlaProLeuThrCysSerProGlnProGlu 430
Db 3529 CCGCTGCCCTGTGAGACTGATGGCTACGTGCCCCCTGAGCTGAGCCCCAGCTGAAA 3588
QY 431 ----- 431
Db 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuPro 450
Db 3589 TATGGAACAGCCAGATGTTCCGGCCAGCCCTTGGCCCGAGAGGCGCTGTGCT 3648
QY 451 ----- 451
Db 451 AlaIleArgProAlaGlyAlaThrLeuGluIleArgAlaLysThrLeuSerProGlyLysAsn 470
Db 3649 GCTGCCGACCTGCTGGTGGCCCACTCGAAGAGGCCCAAGACTCTCTCCCGAGGAAAGAT 3708
QY 471 ----- 471
Db 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
Db 3709 GGGGTGTCAAAGAGCTTTTGGCTTTGGGGTGGCGGTGGAGAACCCCGAATGACTTTGACA 3768
QY 491 ----- 491
Db 491 ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 510
```



```

|||||
Db 3769 CCCGAGGAGAGCTGCCCCCTCAGCCCCCCTCTCTGCTTCAGCCAGCCTTCGAC 3828
Oy 511 AsnLeuTyTrpThrPaspGlnAspProGlnuArgGlyAlaProPserThrPhyls 530
Db 3829 AACCTCTATTACTGTGGACACGAGACCACAGCGGGGGCTCCACCCAGCCTTCAAA 3888
Oy 531 GlyThrProThrAlaGlnAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro 550
Db 3889 GGGACACTAGCGGAGAGAACCCAGACTACTGGCTGTGACGTGCCA-----GTGTGA 3942
Oy 551 AlaArgSerProSerPro 556
Db 3943 ACCGAGAGGCCAAGCTCG 3960

RESULT 15
AAF24297
ID AAF24297 standard; DNA: 9274 BP.
XX
AC AAF24297;
XX
DT 03-APR-2001 (first entry)
XX
DE HER2 transgene plasmid construct.
XX
KW Human; HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
XX antibody; ds.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200100244-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000MO-US17229.
XX
PR 25-JUN-1999; 9905-0141316.
XX
PR 16-MAR-2000; 2000US-0189844.
XX
PA (GETH ) GENENTECH INC.
XX
PI Erickson S, Schwall R;
XX
DR WPI; 2001-061962/07.
XX
DR P-PSDB; AAB60167.
XX
PT Treating tumors, particularly breast cancers, which overexpress an ErbB
PT receptor and does not respond to an anti-ErbB antibody, comprises
PT conjugating the antibody to a maytansinoid -
XX
XX
PS Example 3; Fig 4; 92pp; English.
XX
CC The present invention provides a method of treating cancer by
CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
CC particular, the antibody is directed against ErbB2 (also known as HER2
CC and p185neu). The method is particularly useful in the treatment of
CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.
XX
XX
SO Sequence 9274 BP; 2152 A; 2539 C; 2488 G; 2095 T; 0 other;

Alignment Scores:
Pred. NO.: 3.58e-97 Length: 9274
Score: 2415.00 Matches: 527
Percent Similarity: 41.04% Conservative: 9
Best Local Similarity: 40.35% Mismatches: 32
Query Match: 64.59% Indels: 738
DB: 22 Gaps: 6

US-09-821-883-2 (1-690) x AAF24297 (1-9274)
Oy 2 ArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 21

```

```

|||||
Db 1695 CGCCCCCGCCCTCCAGCCGGGTCCAGCCGAGACCATGAGAGCTGGCG-----GCC 1745
Oy 22 LeuPhePheTrpLeuAspArgSerValLeuAlaLysLeuLeuAlaArgGlyAlaAlaSer 41
Db 1746 TTGTGGCCGTG-----GGGCTCTCTCTGCGCTTGTGGCCCCCGGAGCGGAGC 1796
Oy 42 ThrGlnAlaCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHis 61
Db 1797 ACCCAATGTGACACCGGACAGACATACGTGGGTCTCCCTGCAATGCCGAGACCAC 1896
Oy 62 LeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnLysLeuLeu 81
Db 1857 CTGGACATGCTCCGCCACCTCTACACAGGCTGCGAGGTGTGACAGGAAACCTGAAACTC 1916
Oy 82 ThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnLysValGlnGly 101
Db 1917 ACCTACTGCCCCCAATGACAGCTGTCTCTCTGAGGATATCCAGAGGTGCAGCGGC 1976
Oy 102 TyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuAlaVal 121
Db 1977 TACGTCTCATCGCTCACACCAAGTAGGACAGTCCCACTGCAGAGCGTGGCATTTGTG 2036
Oy 122 ArgGlyThrGlnLeuPheGlnAspSerTyrAlaLeuAlaValLeuAspAsnGlyAspPro 141
Db 2037 CGAGGACCCAGCCTTTGAGGACAACTATGCCCTGCGCTTACGCAATGGAGAGACCG 2096
Oy 142 LeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlyLeuGlnLeu 161
Db 2097 CTGACAAATACACACCCCTGTCACAGGGGCTCCCCAGAGAGCCCTGGGAGCTGCACCT 2156
Oy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeuCys 181
Db 2157 CGAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTATCCAGCGGAACCCCAAGCTGTC 2216
Oy 182 TyrGlnAspThrIleLeuTyrPheLysAspIlePheHisLysAsnGlnLeuAlaLeuThr 201
Db 2217 TACCAAGACACAGATTTGTGGAAGACATCTTCACAAAGAACACCGCTGCTCTACA 2276
Oy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPrometCysLysGlySer 221
Db 2277 CTGATACACACCAACCGCTCTCGGGCTGCCACCCCTGCTCTCGATGTATAGGCTCC 2336
Oy 222 ArgCysTrpGlyLysSerSerGlnAspCysGlnSerLeuThrArgThrValCysAlaGly 241
Db 2337 CGCTGCTGGGGAGAGAGTCTGAGGATTTGACAGCCTGACGGCAGCTGTGTGGCGGT 2396
Oy 242 GlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlyAlaAla 261
Db 2397 GGGTGTCCCGCTGCAGAGGGGCCACTGCCACATGACTGTGCTGCCATGAGAGTGTGCC 2456
Oy 262 GlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly 281
Db 2457 GGGCGACAGGGGCCCAAGCACTGACTGCTGCGCTGCCCTCCACATCAACACAGAGGGC 2516
Oy 282 IleCysIleLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSerMet 301
Db 2517 ATCTGTAGCTGACCTGCCAGCCCTGTGACCTACACACAGACAGCTTTGTAGTGCATG 2576
Oy 302 ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn 321
Db 2577 CCCAATCCGAGGGCCGATTTACATTGGCGCCAGCTGTGTGACTGCTCTCTCCCTACAC 2636
Oy 322 TyrLeuSerThrAspValGlySer----- 329
Db 2637 TACCTTTCTAGGAGCGTGGATCTGTGACACCTGTGTGCCCTGACACCAAGAGAGTG 2696
Oy 329 ----- 329
Db 2697 ACAGACAGAGATGAGAACACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCGAGTGTGC 2756
Oy 329 ----- 329

```


D 2157 TATGCTGGGCAATGAGCACTGCGAGAGTGAGGGCAATTACAGTGCATATATCAG 2816
OY 329 ----- 329
D 2817 GAGTTTGCTGGCTGCAAGAAATCTTTGGAGCCTGGCAATTTCTGCCGAGAGCTTTGAT 2876
OY 329 ----- 329
D 2877 GGGAGCCAGCCTTCACACACTGCCCCGCTCCAGCCAGAGAGCTCAAGTGTAGACT 2936
OY 329 ----- 329
D 2937 CTGGAAGATCACAGGTTACTATACATCTCAGCATGGCCGAGACCTGCTGACCTC 2996
OY 329 ----- 329
D 2997 AGCGCTTCCAGAACTGCAAGTATCCGGGAGCAATTTGCACAATGGCGCTACTCG 3056
OY 329 ----- 329
D 3057 CTGACCCCTGCAGAGGCTGGGCAATCAGCTGGGCTGGGCTCAGAGGGAATGGGC 3116
OY 329 ----- 329
D 3117 AGTGGACTGGCCCTCATCACATACACCCACTCTGCTGCTGCAAGGTCCTCG 3176
OY 329 ----- 329
D 3177 GACCACTCTTTGGAAACCCGACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGGAC 3236
OY 329 ----- 329
D 3237 GAGTGTGGGCGAGGGCTGGGCTGCCACAGCTGTGCCGCCAGGCACTGTGGGT 3296
OY 329 ----- 329
D 3297 CCAAGGCCCAACCACTGTCTCACTGACGCAATTCCTTGGGGCCAGAGTCCGTGAG 3356
OY 329 ----- 329
D 3357 GAATGCCAGTACTGACAGGGGCTCCCAAGGAGTATGTGAATGCCAGGCACTGTTGCCG 3416
OY 329 ----- 329
D 3417 TGCCACCTGAGTGTACGCCCAAGATGGCTCAGTGACTGTTTGGACCGGAGGCTGAC 3476
OY 329 ----- 329
D 3477 CAGTGTGGCTGTGCCCACTATAGAGACCTTCCTTGCCTGGCCGCTGCCAGC 3536
OY 329 ----- 329
D 3537 GGTGTAAACCTGACCTCTCTACATGCCATCTGGAAGTTTCCAGATGAGAGGCGCA 3596
OY 329 ----- 329
D 3597 TGGCAGGCTTGGCCCATCACTGCAACCCACTCTGTGTGACCTGATGACAAAGGCTGC 3656
OY 329 ----- 329
D 3657 CCCGCGAGAGAGAGCCAGCCCTGACGTCATCGCTCTGCGGTGGATTCTG 3716
OY 329 ----- 329
D 3717 CTGCTGTGCTTTGGGGGTGGCTTTGGGATCTCTCATCAAGCAGCGACAGACATC 3776
OY 329 ----- 329
D 3777 CGAAGTACAGATCGGAGACCTGTGAGGAACGAGAGCTGTGGAGCCGCTGACACT 3836
OY 329 ----- 329
D 3837 AGCGAGCGATGCCCAACGAGCGAGATGCGATCTGAAAGAGAGGAGCTGAGAG 3896

OY 329 ----- 329
D 3897 GTGAAGTGTGATGCTGGGCTTTTGGCACACTTACAAAGGCACTGATCCCTGAT 3956
OY 329 ----- 329
D 3957 GGGAGAAATGTAATAATTCAGTGGCCATCAAGTGTGAGGAAACACATCCCCAA 4016
OY 329 ----- 329
D 4017 GCCAACAAGAAATCTTAGACGAAGCATACGTATGCTGTGGCTCCCATATGTC 4076
OY 329 ----- 329
D 4077 TCCGCTTGTGGCATCTGCTGACATCCAGGTGCAGCTGTGACACACTTATGCC 4136
OY 329 ----- 329
D 4137 TATGCTGCTCTTAGACATGTCCGGGAAACCGGAGGCTGGGCTCCAGACCTG 4196
OY 329 ----- 329
D 4197 CTGAAGTGTATGAGATTCACAGGGATGAGCTACCTGAGAGATGTGCGCTGTA 4256
OY 329 ----- 329
D 4257 CACAGGACTTGGCGCTCGAAGCTGCTGTCAGAGTCCCAACCATGTCAAAATTACA 4316
OY 329 ----- 329
D 4317 GACTTGGGCTGGCTCGCTGCTGACATTGACGAGACAGATACATGAGATGGGGC 4376
OY 329 ----- 329
D 4377 AAGTGCCATCAAGTGAATGCGCTGAGTCCATCTCCGCGCGCTTACACCACAG 4436
OY 329 ----- 329
D 4437 AGTATGTGTGAATATGATGTGCTGTGTGAGAGCTGATGACTTTGGGGCCAAACT 4496
OY 329 ----- 329
D 4497 TACATGAGATCCAGCCCGGAGATCCCTGACCTGTGMAAAGGAGAGCGCTGCC 4556
OY 329 ----- 329
D 4557 CAGCCCCATCTGACACATGATGTACATGATCAATGTGATGATTTGAC 4616
OY 329 ----- 329
D 4617 TCTGAATGTGGCCCAAGATTCCGGAGTTGTGTCTGAATTTCCCGCATGGCAGGAC 4676
OY 329 ----- 329
D 4677 CCCAGGCTTTGTGTATCAGATGAGACTTGGGCCAGCAGTCCCTTGACAGC 4736
OY 329 ----- 329
D 4737 ACCTTACCGCTCACTGTGAGAGAGATGACATGGGGACCTGTGATGCTGAGAG 4796
OY 330 ----- 334
D 4797 TATCTGATCCCGACGAGGCTTTCTGTGTCACACCTGCCCGGCGCTGGGGGATG 4856
OY 335 ValHisArgHisArgSerSerThrArgSerGlyGlyIysPleuThrLeuGly 354
D 4857 GTCCACCAAGGACCCGACCTCATCAACGAGGTGGCGGTGGGACTACACTAGGG 4916
OY 355 LeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGluGlyAlaGly 374
D 4917 CTGAGCCCTCTGAAAGAGAGGCCCCAGCTCTCAGTCCAGACCTCCGAAGGGGCTGGC 4976

```
OY 375 SerAspValPheaspGlyaspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuPro 394
      |||||||
Db 4977 TCCGATGTATTTGATGTGACCTGGGAATGGGGGACCAAGGGGCTGCCAAAGCCTCCCC 5036
OY 395 ThrHisApproSerProLeuGlnArgTyrSerGlnAspproThrValProLeuProSer 414
      |||||||
Db 5037 ACACATAGCCCGACCCCTCTACAGGGGTACAGTACAGAGACCCACACATACCCCTGCTCT 5096
OY 415 GluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyrValAsnGln 434
      |||||||
Db 5097 GAGACGTAGTGGCTACGTGGCCCCCTGACCTGCAGGCCCCAGCCTGATATGTGAACACAG 5156
OY 435 ProAspValArgProGlnProProSerProArgGluGlyProLeuProAlaAlaArgPro 454
      |||||||
Db 5157 CCAAGATGTGGGGCCCGCCCTTCGCCCCGAGAGGGCCCTCTGCTGCTGCCGACT 5216
OY 455 AlaGlyAlaThrLeuGlnuArgAlaLysThrLeuSerProGlyLysAsnGlyValLys 474
      |||||||
Db 5217 GCTGGTCCACTCTGGAAAGGCCAAGACTCTCTCCCGAGGAGAAATGGGGTGTCAAA 5276
OY 475 AspValPheAlaPheGlyGlyAlaValAlaGlnAsnProGluTyrLeuThrProGlnGly 494
      |||||||
Db 5377 GACGTTTTGCTTTGGGGGTGGCGTGAGAACCCGAGTACTTGACACCCCGAGGAGA 5336
OY 495 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 514
      |||||||
Db 5337 GCTGCCCCCTGAGCCCGCCCTCTCTGCTTCAGCCAGCCTTCGACAACTCTATTAC 5396
OY 515 TrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGlyThrProThr 534
      |||||||
Db 5397 TGGGACCAAGACCCACAGAGCGGGGGCTCCACCCAGACCTTCANAGGACACCTACG 5456
OY 535 AlaGlnAsnProGluTyrLeuGlyLeuAspValProAlaAlaProAlaArgSerPro 554
      |||||||
Db 5457 GCAGAGAACCCAGAGTACTGGGTCTGCAGTGCCA-----GTGTGACCAAGAGGCA 5510
OY 555 SerProSerThrGlnProThrGlnHisValAsnAlaIleGlnGluAlaArgArgLeu 574
      |||||||
Db 5511 AGTCCGCGAG---AAGCCCTGATGTGTCTCAGG-----GAGCAGGAGAGGGGCTCTG 5561
OY 575 AsnLeuSerArgAspThr 580
      |||||||
Db 5562 AGCTATTCAGAGTACT 5579
```

Search completed: May 8, 2003, 13:04:15
Job time : 1665 secs

GenCore version 5.1.4.p5.4576
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 8, 2003, 10:54:14 ; Search time 4337 Seconds

(without alignments)
4630.144 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRNAPLLLRASLSIGFLF.....EPVQSGAPRRRAHHNNHH 690

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=alp
-Q/cg2_1/USFTO.spool/US09821883/runat_28042003_104749_14566/app_query.fasta_1.839
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09821883 -GCN 1.1 3745 -runat_28042003_104749_14566 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEVTIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg_hum:*
31: em.htg_inv:*
32: em.htg_other:*
33: em.htg_mus:*
34: em.htg_pln:*
35: em.htg_rod:*
36: em.htg_mam:*
37: em.htg_vrt:*
38: em.sy:*
39: em.htgo_hum:*
40: em.htgo_mus:*
41: em.htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3739	100.0	2070	6 AX268286	AX268286 Sequence
2	3473.5	92.9	2091	6 AX268288	AX268288 Sequence
3	2974	79.5	1665	6 AX268285	AX268285 Sequence
4	2959.5	79.2	1692	6 AX268287	AX268287 Sequence
5	2423	64.8	4530	6 AR202597	AR202597 Sequence
6	2423	64.8	4530	6 AX282577	AX282577 Sequence
7	2423	64.8	4530	6 BD005474	BD005474 Cellular
8	2423	64.8	4530	6 121124	121124 Sequence 9
9	2423	64.8	4530	6 159745	159745 Sequence 9
10	2423	64.8	4530	6 HUMHER2A	M11730 Human tyros
11	2422	64.8	9274	6 AX060703	AX060703 Sequence
12	2418	64.7	4473	6 AR080259	AR080259 Sequence
13	2418	64.7	4473	6 AR167390	AR167390 Sequence
14	2418	64.7	4473	6 HSEBR2R	X33363 Human c-erb
15	2410.5	64.5	3768	6 AX060704	AX060704 Sequence
16	2410.5	64.5	3768	6 AX467229	AX467229 Sequence
17	2410.5	64.5	3768	6 AX481438	AX481438 Sequence
18	2405.5	64.3	3768	6 AR034479	AR034479 Sequence
19	2405.5	64.3	3768	6 AX201817	AX201817 Sequence
20	2405.5	64.3	3768	6 AX380923	AX380923 Sequence
21	2405.5	64.3	3768	6 AX384604	AX384604 Sequence
22	2405.5	64.3	3768	6 AX465456	AX465456 Sequence
23	2258.5	60.4	2385	6 AR082744	AR082744 Sequence
24	2258.5	60.4	2385	6 AR099633	AR099633 Sequence
25	2258.5	60.4	2385	6 AR143949	AR143949 Sequence
26	2188.5	58.5	1437	6 AX268289	AX268289 Sequence
27	2130	57.0	2781	6 AX380944	AX380944 Sequence
28	2129	56.9	2763	6 AX380942	AX380942 Sequence
29	2124.5	56.8	3780	4 AB008451	AB008451 Carls lam
30	1968	52.6	3955	6 AX380924	AX380924 Sequence
31	1968	52.6	3955	6 121129	121129 Sequence 14
32	1968	52.6	3955	6 159750	159750 Sequence 14
33	1968	52.6	3955	10 RNNEUR	X03362 Rat mRNA fo
34	1961	52.4	3771	6 AX189662	AX189662 Sequence
35	1961	52.4	3771	6 AX380925	AX380925 Sequence
36	1955	52.3	4062	10 HANNEU	D16285 Syrian gold
37	1680.5	44.9	2871	6 AX023363	AX023363 Sequence
38	1647	44.0	1872	6 AR181000	AR181000 Sequence
39	1613	43.1	1316	9 AF177761	AF177761 Homo sapi
40	1587	42.4	867	6 AX268303	AX268303 Sequence
41	1199.5	32.1	1191	6 AX268307	AX268307 Sequence
42	1188	31.8	1806	6 AX384608	AX384608 Sequence
43	1183	31.6	1755	6 AX384609	AX384609 Sequence
44	1183	31.6	1767	6 AX384607	AX384607 Sequence
45	1183	31.6	1773	6 AX384610	AX384610 Sequence

RESULT 1

ALIGNMENTS

AX268286
LOCUS AX268286 2070 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 7 from Patent WO01/4855.
ACCESSION AX268286
VERSION AX268286.1 GI:16541538
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Laus, R., Vidovic, D. and Graddis, T.
TITLE Compositions and methods for dendritic cell-based immunotherapy
JOURNAL Patent: WO 01/4855-A 7 11-OCT-2001;
DENDREON CORPORATION (US)
FEATURES
source 1. 2070
Location/Qualifiers
BASE COUNT 425 a 707 c 568 g 370 t
ORIGIN
Alignment Scores:
Pred. No.: 5,18e-129 Length: 2070
Score: 3739.00 Matches: 690
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-821-883-2 (1-690) x AX268286 (1-2070)
QY 1 MetArgAlaAlaProLeuLeuLeuAlaAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
DB 1 ATGAGAGCTGACACCCCTCTCTGCGACGAGCAAGCTTAGCTTGCTTCTTGT 60
QY 21 LeuLeuPhePhePheLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
DB 61 CTGCTTTTCTTCTGCTGAGACCGAAGTGTACTAGCCAAAGATTGGCGCGCGCGCG 120
QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCGAGCCCAAGTGTGCACGCGGACAGACATGACGTGGCTCTCTGCCAGTCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuTyrgLgCysGlnValValGlnGlyAsnLeuGlu 80
DB 181 CACCTGGAATGCTGCCGCCCTCTACAGAGGCTGCGAGGTGGCGAGGAACCTGAA 240
QY 81 LeuThrTyrlleuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
DB 241 CTGACCTACCTGCGCCCAATGCGAGCTCTCTCTGCGAGTATCCAGAGGTGCAG 300
QY 101 GlyTyrgValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnAlaArgIle 120
DB 301 GGGTACGTGCTCATCGCTCAACAACCAAGTAGGAGGTCCACAGTCCAGAGGCTCGGATT 360
QY 121 ValArgGlyThrGlnLeuPheGlnAspAsnTyrlAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 361 GTGCGAGGACCCAGCTCTTGAGGACAACTATGCTGCGCTGCTGAGCAATGAGAC 420
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuAspGlyLeuGln 160
DB 421 CCGGTGAACAATACACCCCTGTACAGGGGCTCCCGAGAGAGGCTGCGGAGCTGCAG 480
QY 161 LeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180
DB 481 CTTCGAAGCCTCAAGAGATCTTGAAAGAGGGGTCTTATCCAGGCGGAACCCCACTC 540
QY 181 CysTyrgLlnAspThrIleLeuTyrlPlysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
DB 541 TGTCTACAGAGACAGATTTTGTGGAAGACATCTTCCACAAGAAACACGCTGCTCTC 600
QY 201 ThrGlnIleAspThrAsnAlaSerArgAlaCysHisPheProCysSerProMetCysLysGly 220

DB 601 ACACTGATGAGACACCAACCGCTCTGCGGCTGCGCACCCCTGTCTCTCGAGTGTAAAGGC 660
QY 221 SerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240
DB 661 TCCCGCTGCTGCGGAGAGAGATTGTGAGAGATTGTACAGACCTGACGCCACTGTCTGTGCC 720
QY 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlnGlnCysAla 260
DB 721 GGTGCTGTGCGGCTGCAAGAGGGCCACTGCGCCACTGACTGTCTGATGAGACATGTGCT 780
QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
DB 781 GCCGCTGACAGGGGCCCAAGCACTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 281 GlyIleCysGlnLeuHisCysProAlaLeuValThrTyrgAsnThrAspThrPheGlnSer 300
DB 841 GGCATCTGTGAGCTGACACTGCGCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 301 MetProAsnProGluGlyArgTyrlThrPheGlyAlaSerCysValThrAlaCysProTyrl 320
DB 901 ATGCCCAATCCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGACTGCTGCTGCTGCTGCT 960
QY 321 AsnTyrlleuSerThrAspValGlySerGlyValaGlyIleMetValHisArgHisArg 340
DB 961 AACTACTTCTTCTACGAGCTGCGGATCGCGGCTGCGGCGATGGTCCACACAGCCAGCCGCC 1020
QY 341 SerSerSerThrArgSerSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
DB 1021 AGCTCATCTACAGAGATGGCGGTGGGAGCTGACACTAGAGGCTGAGAGCCCTCGAAGAG 1080
QY 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
DB 1081 GAGCCCCCAGAGTCTCCACTGCGCACCTCCGAAAGGGCTGGCTCGATGTATTTGAAGGT 1140
QY 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
DB 1141 GACCTGGAAATGGGGGAGCGCAAGGGCTCGAAGCCCTCCCGACACATGACCCAGCCCT 1200
QY 401 LeuGlnArgTyrlSerGlnAspProThrValProLeuProSerGluThrAspGlyTyrlVal 420
DB 1201 CTACAGGCGGTACAGTAGAGAGCCCAAGTACCTGCTGCTGAGTATGCTGAGTGT 1260
QY 421 AlaProLeuThrCysSerProGlnProGluTyrlValAsnGlnProAspValArgProGln 440
DB 1261 GCCCCCTGACCTGACCCCGCCAGCTGAATATGTGAACCAAGCAGATGTTGGCCCCAG 1320
QY 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyValThrLeuGlu 460
DB 1321 CCCCCTTCCGCCCGAGAGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
DB 1381 AGGGCCAAAGACTCTCTCCAGAGGAAGATGGGCTCTCAAAAGACGTTTTTGGCTTTGGG 1440
QY 481 GlyAlaValGlnAsnProGluTyrlLeuThrProGlnGlyAlaAlaProGlnProHis 500
DB 1441 GGTGCTGTGAGAAACCCCGAGTACTTGACACCCAGGAGAGAGTGTGCTCCAGCCCCAC 1500
QY 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrlTyrlProAspGlnAspProPro 520
DB 1501 CTTCCTCTGCTGCTGACCCAGCTTGCACACTCTATTAAGTGGAGCAAGAGAGAGAGAG 1560
QY 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGlnAsnProGluTyrl 540
DB 1561 GAGGAGGGGCTGCACACCACTTCAAGAGGACACTGAGGAGAGAGAGAGAGAGAGAGAG 1620
QY 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSerThrGlnPro 560
DB 1621 CTGGGTCTGAGAGCTGCGAGGGCGGACCCGCGCTGCGCCAGCCACACACAGCCCC 1680
QY 561 TrpGlnHisValAsnAlaIleGlnGluAlaArgArgLeuAsnLeuSerArgAspThr 580

Db 1681 TGGGAGCATGTGAATGCCATCCAGAGGCCGCGTCTTCGTAACCTGATGAGACACT 1740
 QY 581 AIAAIAAGIUMETASNGIUNTHRVAlGIVAlIISerGIUMETPhEASpLEuNGIUNPro 600
 Db 1741 GCTGCTGAGATGAATGAACAGTACAGTATCTCAGAAATGTTTGACCTCCAGAGGCCG 1800
 QY 601 ThrcysLeuGIUNThrcysLeuGIUNThrcysLeuGIUNThrcysLeuGIUNThrcysLeu 620
 Db 1801 ACCGCTACACAGACCGCGCTGAGCTGTACAGAGCGGCTCGGGGAGCGCTCACACAG 1860
 QY 621 LeuLYSGIUNProLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeu 640
 Db 1861 CTCAAGGGCCCTTACACCTGATGAGCCACCACTACAAACAGACCTGCGCTCAACCCCG 1920
 QY 641 GIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeu 660
 Db 1921 GAAACTCTCTGTCACACCCAGATTTACACTTTGAAAGTTTAAAGAGAACTGAAAGGAC 1980
 QY 661 PhEASpLEuValIISerPhEASpLEuValIISerPhEASpLEuValIISerPhEASpLEu 680
 Db 1981 TTTCTGCTGTATCCCTTGTACTGCTGTGAGCCAGCTCCAGAGAGGGCGCGCACCCCG 2040
 QY 681 PROAlAIA 690
 Db 2041 CCGGCGGCGCGCACATCACCATCACATCAC 2070
 RESULT 2
 AX268288
 LOCUS AX268288 2091 bp DNA linear PAT 29-OCT-2001
 DEFINITION Sequence 9 from Patent WO0174855.
 ACCESSION AX268288
 VERSION AX268288.1 GI:16541540
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 AUTHORS Laus, R., Vidovic, D. and Gradis, T.
 TITLE Compositions and methods for dendritic cell-based immunotherapy
 JOURNAL Patent: WO 0174855-A 9 11-OCT-2001;
 DENDREON CORPORATION (US)
 FEATURES
 source location/Qualifiers
 1..2091
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="HER500*-rhm-csf construct"
 BASE COUNT 447 a 699 c 559 g 386 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,696-119 Length: 2091
 Score: 3473.50 Matches: 641
 Percent Similarity: 94.428 Conservative: 19
 Best Local Similarity: 91.708 Mismatches: 28
 Query Match: 92.90% Indels: 11
 DB: 6 Gaps: 2
 US-09-821-883-2 (1-690) x AX268288 (1-2091)
 QY 1 MetaGAlAIA 20
 Db 1 ATGAGAGCTGACACCTCTCTCTGCGCAGGAGCAGCAAGCTTAGCCTTGCTCTCTGTTT 60
 QY 21 LeuIUNPhEASpLEuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeu 40
 Db 61 CTGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 41 SerThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 60
 Db 121 TCGACCAAGTGTGACCGGCGACAGCAAGAGCTGCGCTCCCTCCAGCTCCCGAGACC 180
 QY 61 HisLeuAspMetLeuArgHisLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeu 80

Db 181 CACTGACATGCTCCGCCACCTTACCAGGGCTCCAGAGTGTGACAGGAAACCTGGAA 240
 QY 81 LeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 100
 Db 241 CTACACTTACCTCCCAACCAATCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
 QY 101 GIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 120
 Db 301 GCGTACGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 QY 121 ValArgGIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 140
 Db 361 GTCCAGAGCACCACCTCTTGTAGACAACTATGCTGCGCTGCTGCTGCTGCTGCTGCTG 420
 QY 141 ProLeuAspAspThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 160
 Db 421 CCGGTGAACAAATACACCCCTCTACAGGGGCTCCCAAGAGGCTCGGGAGCTGACG 480
 QY 161 LeuArgSerLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 180
 Db 481 CTTCAGAGCTCACAGAGATCTTGAAGAGAGGGCTTGTATCCAGCGGAAACCCAGAGCTC 540
 QY 181 CysTYRGIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 200
 Db 541 TCTTACACAGACACGATTTTGTGAAGGACATCTTCCACAAACACACCTGCTGCTC 600
 QY 201 ThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 220
 Db 601 ACACGTGATGACACACACCGCTCTGCGGCTCCACCCCTGTTCTCCGATGTGTAAGGCG 660
 QY 221 SerArgcysTYRGIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 240
 Db 661 TCCCGCTGCTGGGAGAGAGTCTGAGATTCTCAGAGCTGACAGCGGCTGCTGCTGCTG 720
 QY 241 GIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 260
 Db 721 GGTGCTGCTGCGCGCTGCAAGGGCGCACGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 261 AIAAGIcysTYRGIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 280
 Db 781 GCCGCTGACAGCGGCCCAAGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 281 GIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 300
 Db 841 GGCATCTGAGCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 900
 QY 301 MetProAspProGIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 320
 Db 901 ATGCCCAATCCCGAGGGCGGTATACATTCGGCGCAGCTGTGACTGCTGCTGCTGCTGCT 960
 QY 321 AsnTYRLeuSerThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 340
 Db 961 AACCTACCTTTCTACGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1020
 QY 341 GIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 360
 Db 1021 GGGGGCATGCTCACACAGAGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 QY 361 ThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 380
 Db 1081 ACACCTAGAGCTGAGCGCTTGAAGAGAGGCGCCAGGCTGCTGCTGCTGCTGCTGCTGCT 1140
 QY 371 GIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 390
 Db 1141 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
 QY 391 SerLeuProThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 410
 Db 1201 ACCCTCCCAACACAGACCCAGCTTACAGCGGTACAGAGAGACCCACACATATACC 1260
 QY 411 LeuProSerGIUNThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcysLeuThrcys 430
 Db 1261 CTGCTCTGAGACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

QY 432 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 451
 DB 1331 GTGAACACGAGCATGTTGGGCGCCAGCCCTTGCCCGAGAGAGGGCCCTGCTGCT 1380
 QY 452 AlaArgProAlaGlyAlaThrLeuGlnArgAlaLysThrLeuSerProGlyLysAsnGly 471
 DB 1391 GCGCCACCTGGTGGTCCACCTCTGMAAGGGCCAAAGACTCTCTCCCGAGGAGAAATGG 1440
 QY 472 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluThrLeuPro 491
 DB 1441 GTCGCAAAAGACGTTTGGCTTTGGGGGTGGCGTGGAGAACCCGAGTACTTGACACC 1500
 QY 492 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 511
 DB 1501 CAGGAGAGAGCTGCGCCCTCAGCCGACCCCTCTCTGCTTCAAGCCGCTTGACACAC 1560
 QY 512 LeuThrThrThrPheArgLysAspProGluArgGlyAlaProProSerThrPheLysGly 531
 DB 1561 CTCTATTTCTGGGACAGAGACCCACAGAGGGGGGCTCCACCCAGACCTTCMAAGGG 1620
 QY 532 ThrProThrAlaGluAsnProGluThrLeuGlyLeuAspValProAlaAlaAlaProAla 551
 DB 1621 ACACCTAGGCGAGAAACCCAGAGACTGGGTCTGGACGTGCCAGCGCGCCGCCACC 1680
 QY 552 ArgSerProSerProSerThrGlnProThrPheLysValAsnAlaLysGlnLysAlaArg 571
 DB 1681 CGCTACACCAACCTGCTCAGCCGCGCGCTGAGAACATGTGATGCCATCAAGAAAGCTCTG 1740
 QY 572 ArgLeuLysAsnLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValLe 591
 DB 1741 AGCCCTCAATATGACATCGCTCTGTGGAGAACGMAAGAGAGAGATGACATCATCMT 1800
 QY 592 SerGlnMetPheAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuThrLys 611
 DB 1801 TCTATATAGTTCTCATCATCAGAGCGGACATGTGTGACAGCCCGCTTAACCTATACAG 1860
 QY 612 GlnGlyLeuArgGlySerLeuThrLysLeuLysGlyProLeuThrMetMetAlaSerHis 631
 DB 1861 CAGGCTACAGGGGCGACCTCAACCAACTCAATGAGCGCTTGACATGATGACGAC 1920
 QY 632 TyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGlnIleLeuThrPhe 651
 DB 1921 TACCAAGCAACTGCCCTCAACCCCGGAACTGACTGAATGAAGTACACCACTTT 1980
 QY 652 GluSerPheLysGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGlu 671
 DB 1981 GAGGATTTCAATAAAGAACTTAAGGCTTTCTGTTGATATCCCTTTGATGCTGCGAAG 2040
 QY 672 ProValGlnGluGlyAlaProProProProAlaAlaAlaHisHisHisHisHis 690
 DB 2041 CCGGTCCAGAAAGGCGCCACCCCGCGCG-----GCGCATCAACATCACATCAC 2091
 RESULT 3
 AX268285 1665 bp DNA linear PAT 29-OCT-2001
 LOCUS AX268285
 DEFINITION Sequence 6 from Patent WO0174855.
 ACCESSION AX268285
 VERSION AX268285.1 GI:16541537
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM artificial sequences.
 REFERENCE
 AUTHORS Lous R., Vidovic D. and Graddis T.
 TITLE Compositions and methods for dendritic cell-based immunotherapy
 JOURNAL Patent: WO 0174855-A 6 11-OCT-2001;
 DENDREON CORPORATION (US)
 FEATURES
 source 1. 1665
 location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="HER500 construct"

BASE COUNT 330 a 568 c 465 g 302 t
 ORIGIN
 Alignment Scores:
 Score: 4.05e-101 Length: 1665
 Pred. No.: 2974.00 Matches: 554
 Percent Similarity: 99.46% Conservative: 0
 Best Local Similarity: 99.46% Mismatches: 1
 Query Match: 79.54% Indels: 3
 DB: 6 Gaps: 1
 US-09-821-883-2 (1-690) x AX268285 (1-1665)
 QY 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
 DB 1 ATGAGAGCTGACCCCTCTCTCTGCGAGGCGACAGCAACCTTACCTTGGCTTGTGTT 60
 QY 21 LeuLeuPhePheThrPheAspArgSerValLeuAlaLysGlyLeuAlaArgGlyAlaAla 40
 DB 61 CTGCTTTTCTGTGGCTAGACCGAAGTGTACTAGCCAAAGAGTTGGCGCGGGGCGCG 120
 QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
 DB 121 TCGACCCAGTGTGACCGGCGACAGACTGAAGCTGGGCTCTCTGCGAGTCCGAGAC 180
 QY 61 HisLeuAspMetLeuArgHisLeuThrGlnLysCysGlnValValGlnLysAsnLeuGlu 80
 DB 181 CACCTGAGCATGTCTCCGCGACCTCTACACAGGGGTGCCAGGTGGTGCAGAGAACTGGAA 240
 QY 81 LeuThrThrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
 DB 241 CTCAGTACCTGCGCCACCAATGCCAGCTGTCTTCCGAGATATCCAGAGAGTGCAG 300
 QY 101 GlyThrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgLys 120
 DB 301 GCGTACGTCGTATCTCCCTCAACCAAGTGAAGGAGGTCCACGTGCAGAGCTCGGAT 360
 QY 121 ValArgGlyThrGlnLeuPheGluAspAsnThrAlaLeuAlaValLeuAspAsnGlyAsp 140
 DB 361 GTGGAGAGCACCCAGCTCTTGAAGACAACTATGCCCTGCGCGGTGATCAATAGGAGAC 420
 QY 141 ProLeuAsnAsnThrThrProValThrGlyLysAspProGluGlyLeuArgGluLeuGln 160
 DB 421 CCGGTGAACAATACCAACCCCTGTCAAGGGGCTCCCGAGAGGCTCGCGGACTGTGAG 480
 QY 161 LeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAspProGlnLeu 180
 DB 481 CTTCGAACCTCACAGAGATTTGAAGAGAGGGCTTGAATCCAGCGGAACCCAGCTC 540
 QY 181 CysThrGlnAspThrIleLeuThrPheLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
 DB 541 TGTACACAGGACAGATTTGTGGAGAGACATCTTCCACAAGAACCAACACCTGGCTCTC 600
 QY 201 ThrLeuIleAspThrAspArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
 DB 601 ACCTGATAGACACCAACCGCTCTGGGCTGCCACCCCTCTCTCCGATGTGAAGGCG 660
 QY 221 SerArgCysTrpGlyGluSerSerGlnAspCysGlnSerLeuThrArgThrValCysAla 240
 DB 661 TCCGCTCTCTGGGAGAGAGTTCTGAGGATTTGTCAAGCCTGACGCGCAGCTGTGTGCC 720
 QY 241 GlyLysCysAlaArgCysLysGlyProLeuProThrAspCysHisLysGlnCysAla 260
 DB 721 GGTGGCTGTGCCCTGCCAAGGGGCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
 DB 781 GCGGCTGTCAGAGGCGCCCAACACTGTACGCTGTGGCTGTCCCTCACTTAACACAGT 840
 QY 281 GlyLysCysGluLeuHisCysProAlaLeuValThrThrAsnThrAspThrPheGluSer 300
 DB 841 GGCATCTGTAGCTGCACCTGCCAGCCCTGTGTACCTACACACAGACACGTTTGAGTCC 900

```

Oy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db 901 ATGCCCAATCCGAGAGGGCGGATACATTGGCGCCAGCTGTGTACAGCCGTCCCTAC 960
Oy 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArg 340
Db 961 AACTACCTTTTACACGACGTGGAGATCGGCGCTGGGGCATGTGTCCACACAGCACCGC 1020
Oy 341 SerSerSerThrArgSerGlyGlyValAspLeuThrLeuGlyLeuGluProSerGluGlu 360
Db 1021 AGCTATCTACAGAGATGGCGGTGGGACCTGACACTAGGGCTTGAGCCCTCTAAGAG 1080
Oy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
Db 1081 GAGGCGCCAGGTCTCCACTGGCACCTCCGAGAGGGGTGGCTCCGATGATTTATGTGT 1140
Oy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
Db 1141 GACCTGGGAAATGGGGGACGCCAAGGGGCTGCMAAGCCTCCACACATGACCCAGCCCT 1200
Oy 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
Db 1201 CTACAGCGGTACAGTGAAGACCCACAGATACCCCTGCTGTGACACTGATGGTACGTT 1260
Oy 421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
Db 1261 GCGCCCGCTGACCTGCAGCCCGCAGCTGATATGTGAACACACAGATGTTGGGCCAG 1320
Oy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGln 460
Db 1321 CCCCCCTGGCCCCGAGAGGGCCCTGCTGCTGCCACCCGCTGGGCCCTCTGGAA 1380
Oy 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
Db 1381 AGGCGCAAGACTCTCTCCCGAGAGATGGGTGCTCAAGAGCTTTTGTCTTTGGG 1440
Oy 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHis 500
Db 1441 GGTGCGGTGAGAAACCCGAGTACTTGAACACCCGAGGGAGGAGTCCCTCAGCCCA 1500
Oy 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrThrAspGlnAspProPro 520
Db 1501 CCTCCTCTGCTTACAGCCACCTTCGACACCTCTATTACTGGAGCAGGACCCACCA 1560
Oy 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540
Db 1561 GAGCGGGGGGCTCCACCCAGCACTTCAAGGACACCTACGCGCAGAGAACCCAGAGTAC 1620
Oy 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSer 557
Db 1621 CTGGGCTTGAGAGCTGCAGCGGCC-----GCACA-TCACCATCACATCA 1664

```

```

RESULT 4
AX268287 1692 bp DNA linear PAT 29-OCT-2001
LOCUS Sequence 8 from Patent W00174855.
DEFINITION AX268287
ACCESSION AX268287
VERSION AX268287.1 GI:16541539
KEYWORDS
ORGANISM
SOURCE
synthetic construct.
synthetic construct
artificial sequences.

```

```

REFERENCE
1 Laus, R., Vidovic, D. and Graddis, T.
TITILE Compositions and methods for dendritic cell-based immunotherapy
JOURNAL Patent: WO 0174855-A 8 11-OCT-2001;
DENDREON CORPORATION (US)
location/Qualifiers

```

```

FEATURES
source
1. 1692
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="HER500* construct"

```

```

BASE COUNT 338 a 573 c 470 g 311 t

```

ORIGIN

Alignment Scores:

Pred. No.: 1,4e-100

Score: 2959.50

Percent Similarity: 97.88%

Best Local Similarity: 97.88%

Query Match: 79.15%

DB: 6 Gaps: 2

US-09-821-883-2 (1-690) x AX268287 (1-1692)

```

Oy 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGACACCCCTCTCTCTGCGCAGGAGAGCAACCTTAGCTTGTGTTT 60
Oy 21 LeuLeuPhePheTyrPheAspArgSerValLeuAlaLysGlyLeuAlaArgGlyAlaAla 40
Db 61 CTGCTTTTCTCTGCTGACCGAAGTGTACTAGCCAAAGAGTTGGCGGGGGCCGCG 120
Oy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCAGTGTGCACCGGCACAGACATGAAGCTGGGGTCCCTGCGCAGTCCGAGACC 180
Oy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGln 80
Db 181 CACCTGAGACATGCTCCGCCACCTTACACAGGGCTCCAGTGTGTCCAGGAAACCTGGAA 240
Oy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnIleValGln 100
Db 241 CTCACCTACCTGCGCACCAATCCAGCTGCTTCTGACAGATATCAGAGAGTGCAG 300
Oy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGTACGTGCTCATGTGCTCACAAACAGATGAGCGAGGTCCCACTCAGAGAGTGGGAT 360
Oy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGCACCCACCTCTTTAGAGACAACTATGCCCTGGCCGTCTAGACAAATGGAGAC 420
Oy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGGTGAACAATATCACCAACCCCTGTACAGAGGGGCTCCCGAGAGGCTCCGGAAGCTGCAG 480
Oy 161 LeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAspProIleLeu 180
Db 481 CTTCGAGAGCTTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAACCCCAAGCTC 540
Oy 181 CysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAspAsnGlnLeuAlaLeu 200
Db 541 TGTACACAGACAGATTTTGTGGAAGGACATCTTCCACAAAGAACACAGCTGGCTCTC 600
Oy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACATGATAGACACCAACCGCTCTGCGGCTGSCACCCCTGTTCGATGTGTAAGGGC 660
Oy 221 SerArgCysTyrPheGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValLysAla 240
Db 661 TCCGCTCTGTGGGAGAGATTTGAGATTTGTGAGAGCTGTGAGAGCTGTGAGAGCTGTGAGAGCT 720
Oy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrArgPheCysCysHisGlnIleCysAla 260
Db 721 GGTGCTGTGCGGCTGCGAAGGGGCACTGCCCACTGACTGTGCGATGAGAGAGTGTGCT 780
Oy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCGGCTGACAGGGGCCCCAAGACACTGACTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Oy 281 GlyIleCysGlnLeuHisCysArgProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 900
Oy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320

```

```

DB 901 ATGCCCATCCGAGGCGCGGATATTCATTCGCGCCAGCTGTGTACCTGCTCCCTAC
OY 321 AsnTYrLeuSerThrAspValGlySer-----GlyAla 331
DB 961 AACCTACCTTTTACGAGAGTGGATCCGCTACATCATTAATTTCGAGAGAGTTGGCCCT
OY 332 GlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyLeu 351
DB 1021 GGGGGCATGTGTCACACAGGACCGCAGCTCATCTACAGAGAGTGGGGGAGCTGTG
OY 352 ThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 371
DB 1081 ACACAGAGGCTGAGACCTCTGTAAAGAGAGGCCCCAGGCTTCCACTGACCTCCGAA 1140
OY 372 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 391
DB 1141 GGGGCTGCTCCGATGATTGATGTGTCACCTGGGAATGGGGGAGCCAGAGGCGCTGCA 1200
OY 392 SerLeuProThrHisAspProSerProLeuGlnArgTYrSerGlyAspProThrValPro 411
DB 1201 AGCTTCCACACATGAGCCACGCTCTACAGCGGTACAGTGAAGAGCCACACACTACCC 1260
OY 412 LeuProSerGluThrAspGlyTYrValAlaProLeuThrCysSerProGluProGluTYr 431
DB 1261 CTGCCCTCTGAGACTGATGAGTACCTGCTCCCTGACCTGACGCCCCACAGCTGAATAT 1320
OY 432 ValAsnGlnProAspValArgProGluProProSerProArgGlyGlyProLeuProAla 451
DB 1321 GTGAACCGACGATGTGCGCCCGCCAGCCCTTCCGCCCCAGAGAGGCGCTGCTGCTCT 1380
OY 452 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 471
DB 1381 GCGGACCTGCTGTTGGTCCAGCTGGAAAGGCGCAAGACTCTCTCCCGGAGAAAGATGGG 1440
OY 472 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTYrLeuThrPro 491
DB 1441 GTGCTCAAGAGCTTTTGGCTTTGGGGGTGCGGTGAGAAACCCGAGTACTTGACACCC 1500
OY 492 GlnGlyGlyAlaAlaProGluProHisProProAlaPheSerProAlaPheAspAsn 511
DB 1501 CAGGAGAGAGTGGCCCCCAGCCACCCCTCTCTGCTGACGCCAGCTTGCACAAAC 1560
OY 512 LeuTYrTYrTPAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly 531
DB 1561 CTCGATTTCTGGACACGAGACCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAAGGG 1620
OY 532 ThrProThrAlaGluAsnProGluTYrLeuGlyLeuAspValProAlaAlaAlaProAla 551
DB 1621 ACACCTACGCGCAGAAACCCAGAGTACCTGGTCTGGACGTGCCAGCGGGC-----GCA 1674
OY 552 ArgSerProSerProSer 557
DB 1675 CA-TCACCATCACCATCA 1691

RESULT 5
AR202597 4530 bp DNA linear PAT 20-JUN-2002
LOCUS AR202597
DEFINITION Sequence 4 from patent US 6365151.
ACCESSION AR202597
VERSION AR202597.1 GI:21498768
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 4530)
AUTHORS Halpern, M.S. and England, J.M.
TITLE Cellular immunogens comprising cognate proto-oxogenes
JOURNAL Patent: US 6365151-A 4 02-APR-2002;
FEATURES
Location/Qualifiers
1..4530
/organism="unknown"
BASE COUNT 922 a 1382 c 1346 g 880 t

```

```

ORIGIN
Alignment Scores:
Pred. No.: 1,49e-80 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.148 Conservative: 5
Best Local Similarity: 40.758 Mismatches: 21
Query Match: 64.80% Indels: 736
DB: 6 Gaps: 4

US-09-821-883-2 (1-690) x AR202597 (1-4530)
OY 2 ArgAlaAlaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
DB 94 CGCGCCCTCCCGACCGCGGGTCCAGCGGATGGGCGCGGAGCGGAGTGAACCATG 153
OY 18 PheLeuPheLeuLeuPhePheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
DB 154 GAGCTGGCGGCTTGTGCGGCTGG-----GGGCTCTCTCGCCCTTCCGCCCCC 204
OY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
DB 205 GAGCGCGGAGCACCACCAAGTGTGACCGCGCACAGACATGAAGCTGGGCTCCCTGCCAGT 264
OY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTYrGlnGlyCysGlnValValGlnGly 77
DB 265 CCGGAGACCCACCTGGACATGCTCCGCGCACCTTAACAGGGCTCGACAGTGTGTCAGGGA 324
OY 78 AsnLeuGluLeuThrTYrLeuProThrAsnAlaSerLeuSerPheLeuGluAlaPheGln 97
DB 325 AACCTGGAACTACCTACTCTCCACCAATGCCAGCTGTCTTCTGCAAGATATTCAG 384
OY 98 GluValGlnGlyTYrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
DB 385 GAGGTGACGAGGCTACGTCTCATGCTCACAACCAAGAGGAGGTGCCACTGCAGAGG 444
OY 118 LeuArgGlyLeuValArgGlyThrGlnLeuPheGluAspAsnTYrAlaLeuAlaValLeuAsp 137
DB 445 CTGGGATGTGTGCAAGGACCCAGCTCTTTGAGGACCACTATGGCTGGCGCTCTACAC 509
OY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
DB 505 AATGAGAGCCCGGTGAACAATACCAACCCCTGTACAGAGGGCTCCACAGAGGCTGGG 564
OY 158 GluLeuGlnLeuArgSerLeuThrGluLeuLysGlyGlyValLeuLysGlnArgAsn 177
DB 565 GAGCTGACGCTTCGAAAGCTTCACAGAGATCTTGAAGAGGAGGGCTTGATCCACGGGAC 624
OY 178 ProGlnLeuCysTYrGlnAspThrThrLeuThrPheLysAspLeuPheHisLysAsnGln 197
DB 625 CCCCAGCTCTGCTCCAGAGACGATTTTGTGAAAGGACATCTTCCACAGAGAACCCAG 684
OY 198 LeuAlaLeuThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
DB 685 CTGGCTCTACACATGATAGACACCAACCGCTCTGGGGCTCCACCCCTGTCTCCGATG 744
OY 218 CysLysGlySerArgCysTYrPheGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237
DB 745 TGTAAAGGCTTCCCGCTCTGCGGAGAGATTTCTGAGATTTCTCAGACCTGACCGGACT 804
OY 238 ValCysAlaGlyGlyCysValArgCysValGlyProLeuProThrArgPheCysValGlu 257
DB 805 GTCTGTCCGCGGTGCTGCGGCTGCAAGGGGCGGACCTGCGGCTGCTGCTGCTGCTG 864
OY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
DB 865 CAGTGTCTGCGGCTGACAGGGGCGCCCAAGACATCTGCTGCTGCTGCTGCTGCTGCTG 924
OY 278 AsnHisSerGlyLysCysGlyLeuHisCysProAlaLeuValThrTYrAsnThrAspThr 297
DB 925 AACCAAGAGTGCATCTGTGAGTGTGCTGACGCGGCTGCTGACCTGACCTGACCTGAC 984
OY 298 PheGluSerMetProAsnProGluGlyArgTYrThrPheGlyAlaSerCysValThrAla 317

```


|||||
Db 985 TTGAGTCATGCCCAATCCGAGGGCCGCTATACATTGGCGCCAGCTGTGACTGCC 1044
QY 318 CysProtyrAsnTyrIleuSerThrAspValIglySer----- 329
Db 1045 TGTCCCTCACTACCTTTCTACGGACGTGGGATCTGTCAACCCTGCTGCCCCCTGCAC 1104
QY 329 ----- 329
Db 1105 AACCAAGAGTGCACGACAGAGATGSAACACAGCGGTGTGAGAAAGTCAGCAAGCCCTGT 1164
QY 329 ----- 329
Db 1165 GCCCGAGTGTCTATGTCTGGGCATGAGACTTGCAGAGGTGAGGCGAGTTACAGT 1224
QY 329 ----- 329
Db 1225 GCCAATATCCAGGAGTTTGTGCTGCAGAAAGATCTTTGGAGCCTGGCAATTTCTGCCG 1284
QY 329 ----- 329
Db 1285 GAGAGCTTTGATGGGAGCCAGCCCTCCACACTGCCCGCTCCAGCCAGAGCAGCTCCAA 1344
QY 329 ----- 329
Db 1345 GTGTTGAGACTCTGGAAGAGATCACAGGTACTATACATCTCAGCATGCGCGAGACG 1404
QY 329 ----- 329
Db 1405 GTGCTGACCTGAGCCTCTTCAGAACCTGCAAGTAATCCGGGGAGCAATTTCTGCACAA 1464
QY 329 ----- 329
Db 1465 GGGGCTACTCGCTGACCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTACTG 1524
QY 329 ----- 329
Db 1525 AAGGAACCTGGGCACTGAGACTGGCCCTCATTCACATACACCACCTGTCTGCTGCAC 1584
QY 329 ----- 329
Db 1585 ACGGTGCTGTGGAGACAGCTTTTGGGAACCGCACCAAGCTCTGTCTCCACTGCCAAC 1644
QY 329 ----- 329
Db 1645 CGGCGAGAGAGAGTGTGTGGGCGAGGGCTGGCCTGCCACACAGCTGTGCCCGGAGGG 1704
QY 329 ----- 329
Db 1705 CACTGCTGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCCAGTTCTTGGGGCCAG 1764
QY 329 ----- 329
Db 1765 GAGTGGGTGAGGAATGCCGAGTACTGCAGGGGCTCCAGAGGATATGTGAATGCCAGG 1824
QY 329 ----- 329
Db 1825 CACTGTTGCCGTGCCACCTGAGTGTCAAGCCCAAGATGGCTCACTGACTGTTTGA 1884
QY 329 ----- 329
Db 1885 CCGGAGGCTGACCACTGTGTGGCTGTGCCACTATAGAGACCTCTTCTGCGGTGCC 1944
QY 329 ----- 329
Db 1945 CGCTGCCCAAGGGTGTGAACCTGACTCTCTCTACATGCCCATCTTGGAATTTCCAGAT 2004
QY 329 ----- 329
Db 2005 GAGGAGGGGCAATGCCAGCCTTGCCCCATCACTGCACCCACTCTGTGTGGACTGGAT 2064
QY 329 ----- 329

Db 2065 GACAAAGGCTGCCCGCGAGAGAGACCAAGCCCTGTGACGTCCATCTCTGCGGTG 2124
QY 329 ----- 329
Db 2125 GTTGGCATTTCTGTGTGTGTGTCTTTGGGGTGTGTCTTTGGATCTCTATCAAGCGAGG 2184
QY 329 ----- 329
Db 2185 CAGCAAGAGATCCGGAAATGACAGATGGGAGACTGTCTGACAGAAACGAGCTGGTGAG 2244
QY 329 ----- 329
Db 2245 CCGCTGACACTAGCGGAGCGATGCCAACAGCGCAGATCGGATCTGAAAGAGAG 2304
QY 329 ----- 329
Db 2305 GAGCTGAGAGAGTGAAGTGTCTGTGATCTGGCGCTTTTGGCACAGTCTCAAGGCGATC 2364
QY 329 ----- 329
Db 2365 TGGATCCCTGATGGGAGAAATGTGAATAATTCAGTGGCCATCAAGTGTGAGGGAAAC 2424
QY 329 ----- 329
Db 2425 ACATCCCCCAAGCCAAAGAAATCTTAGAGAAAGATAGTGTGCTGTGTGGGC 2484
QY 329 ----- 329
Db 2485 TCCCATATATGTCTCCGCTCTTGGGCACTGTCTGACATCCACGGTGCAGCTGTGACA 2544
QY 329 ----- 329
Db 2545 CAGCTTATGCCATATGGCTGCCTTTAGACATGTCCGGGAAACCGGGAGCGCTGGGC 2604
QY 329 ----- 329
Db 2605 TCCAGAGACTGTCTGAACTGTGTATGACAGATTGCCAAGGGATGAGTACTGGAGAT 2664
QY 329 ----- 329
Db 2665 GTGCGGCTGTACACAGGACTTGGCCGCTCGGAAGCTGTCTGTCAAGATCCCAACAT 2724
QY 329 ----- 329
Db 2725 GTCAAAATTACAGACTTCGGGCTGGCTGGCTGTGTGACATTGACAGACAGATACAT 2784
QY 329 ----- 329
Db 2785 GCAGATGGGGGCAAGGTGCCATCAAGTGATGGCGTGGATCCATTCTCCGCGCGG 2844
QY 329 ----- 329
Db 2845 TTCACCCACACAGATGATGTGTGAGATTATGGTGTGACTGTGTGGAGCTGATGACTTT 2904
QY 329 ----- 329
Db 2905 GGGGCCAAACTTAGAGTGGATCCAGCCCGGAGATCCTGACCTGCTGGAAGAGGG 2964
QY 329 ----- 329
Db 2965 GAGCGGCTGCCAGCCGCCCATCTGCACCATGTGATGTCTACATGATCATGTGCAATGT 3024
QY 329 ----- 329
Db 3025 TGGATGATTGACTGTGAATGTGTGGCCAAAGATTCGGGAGTGTGTGTGTAATTTCCCGC 3084
QY 329 ----- 329
Db 3085 ATGGCAGGAGCCCGAGCGCTTTGTGTATCATCAAGATAGAGACTTGGGCCAGCCAGT 3144
QY 329 ----- 329
Db 3145 CCCTTGACAGACCTTCTACGCTCACGTGTGAGAGAGATGACATGGGGAGCCTGTGTG 3204

[illegible]

```

CDS
151..3918
/note="unnamed protein product"
/codon_start=1
/protein_18="CAD10597.1"
/db_xref="GI:16609669"
/translataion="MELALALRWGLLALLALPFGASLTVCTGTDMKLRLPASPETHL
MLRLHYQCCVQVQGNLEALTYLPTNASLSFLDDIDVGVYLIANVOYRQVLOLRITR
RGTOLEEDNYALVALNDGDDPLNTPYVAGSLRGLQLRSLREILKGLGLIQRNPF
LCYODTLIMNDLIFPKNNOLDLTLIDTNRSRACHGCSPMKSCRMGESSEDCGLTRT
VACGACARCKGRLPTDCCHEOCACGCTGPGSDCLACPLFHSGLICELHDAITYNNT
DTESPMRPEBRTYFGASCYTRACYNLTSDVSGCTLYPLRNBVTRAEPTQCEKSC
SKPCARVCTGLGMELHLEVRVAVTSANIOERGCKITGSLALPESTDGDPASTAPR
QPELOLVEFTLEITGYLYTISAMPDSLPDLSVEFNLOYIRRLILNANVSLTLQGLG
SMELRLLELRELSGLIALIHNTHLCEFTYVWMDLFRNHLILNTANRPEDCEBEGE
LACIOLARHGCMGPRPQVCANCSQFLRGOCEVECRQLDLPREYVARSLCPHEE
CQPNGSVTCGFPBADOCVACAKKDPFCVARPSGKPLDSMPYMKPFDEGACG
PCPIRNTSCVDLDDKGGPAROASPLSTYSAVGLILVVLGVGGIILKROOKK
KRYMRLLOTELVEYRPTPGAMPNOMNLIKETELRKVKVLSGATGVYVYKIM
PDGENVKIPIAKVLRENTSPKANKELIDENYVNAVGSPIVSRLLGLICTSTVOLYH
QLMYEGCLDHRNENRGLSGQDILNMCMIJAKMSYLEVRVLRDILNANVYLVKSB
NHVITDFGLARLIDIDETEXHADGGVPIKMALESILRRRFTHQSDVSYGVTVBE
LMTGAKPARDIFPAREIPEDLLEKGERLPQPICTIDYMIIVKCMIMDSIECRPFRE
VSESRMARODPVEFVIONEDLPASPLDSEFYSLLEDDMGDLVQAEIEYVPOQEG
FCPCPARGAGSMVHHRRSSSTRSGGDDLTGLIGLSEEAERSLAEEAGASVPFG
DLGMAKAKGDSLTPHPSPLORSDEPTVLYPSEBTVGYVAPLTCSPQEPYVNPVR
PQPSRPEGLPAPARPGATLERRAKTUSPGKNGVGLVFAFGAVENPEYILTPGGAA
PQHPHPAPFSAFENLNYWDDPBERGAPSTFEGKTPAENPEYLGLDVP"
BASE COUNT      922 a 1382 c 1346 g      880 t
ORIGIN
Alignment Scores:
Score: No.:      1.49e-80
Pred. No.:      2423.00
Percent Similarity: 41.14%
Best Local Similarity: 40.75%
Query Match:      64.80%
DB:
      Gaps:      6
              4
US-09-821-883-2 (1-690) x AN282577 (1-4530)
OY      2  ATGAATAAProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
      DB      94  GCGGCCCTCCGACGGGGGTCACGGCGAGACCATGTGGGGCGGAGCGCGACGATGCACCATG 153
OY      18  PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
      DB      154  GAGCTGGGGCGGCTTGTGGCGCGTGG-----GGGCTCTCTCTCGGCCCTTGTGCCCCC 204
OY      38  GLVATAAASerThGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
      DB      205  GGAGCCGGGAGACCAACCAAGTGTGCACCGGCACAGACATGAACTCGGCTCCCTGCCAGT 264
OY      58  ProGluTrhHisLeuAspMetLeuArgHisLysLeuTyrgInglyCysGlnValValjIngly 77
      DB      265  CCCGAGACCACACTGCAGATGCTCCGCCACACCTCTACACAGGGCTGCCAGAGTGCAGAGGA 324
OY      78  AsnLeuGluLeuThrTyTrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
      DB      325  AACCTGGAACTACTACTCTGCGCCACCAATGCGACACCTGTCTCTCTCCAGATATTCAG 384
OY      98  GluValjGInglyTyValLeuIleAlaHisAsnGlnValArgGlnValIProLeuGlnArg 117
      DB      385  GAGGTGCAGAGGCTACGTCTCATGCTCTACACAACCAAGTGAAGAGAGTCCCACTGCAGAGG 444
OY      118  LeuArgIleValArgGlyThGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAsp 137
      DB      445  CTGCGGATTGTGCGAGGACACCAAGCTCTTTTGAGGCAACTATGCGCTGCGCGTGTACAC 504
OY      138  AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
      DB      505  AATGAGACCCGCGCTGAACATACCAACCCCTGTCTACAGAGGGGCTCCCAAGAGAGCTCGGG 564
OY      158  GluLeuGlnLeuArgSerLeuThrclnLeuLeuTyrgInglyValLeuIleGlnArgAsn 177

```

|||||
Db 565 GAGCTGACAGCTTCGAAGCTCAGACAGATCTGAAAGAGGGGCTTGTATCCAGGAGAAC 624
Oy 178 ProGlnLeuCySTyrGlnAspThr1LeuLeuTrrPlyAsp1LePheHisLysAsnAngln 197
Db 625 CCCCAGCTCTGTGTACAGACAGCATTTTGTGGAGGAGCATCTTCACAGAACCAACAG 684
Oy 198 LeuAlaLeuThrLeu1LeAspThrAsnArgSerArgAlaCysHisProCysSerPromet 217
Db 685 CTGGCTCTCACACTATATAGACACACACCGCTCTCGGGCCCTGACACCCCTGTTCTCCGATG 744
Oy 218 CysLysGlySerArgCysTrrPlyGluSerSerGluAspCys6LysLeuThrArgThr 237
Db 745 TGTAAAGGCTCCCGCTGCTGGGAGAGATTCTGAGATTTGACAGCCTGACGGCAGCT 804
Oy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisSglu 257
Db 805 GTCTGTGCGGTGGTGGTGGCCGCTGCAAGGGGCGCACCTGCCACTGCTGCTCCATGAG 864
Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865 CAGTGTGCTGCGGGCTGCACGGGGCCCCCAAGCACCTGTGACTGCTGCTGCTCCACTTC 924
Oy 278 AsnHisSerGly1LeCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
Db 925 AACACAGATGGCATCTGTAGCTGCACCTGCCAGCCCTGCTACCTACACACAGACAG 984
Oy 298 PheGluSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
Db 985 TTTGAGTCCATATGCCCAATCCCGAGGGCGGTATACATTTCGGCGCAGCTGTGTGATGCC 1044
Oy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer 329
Db 1045 TGTCTCTACAACTACCTTTCTAGACAGCTGGATCTCTGACCCCTGCTGCCCTGCAC 1104
Oy 329 ----- 329
Db 1105 AACCAAGAGTGCACAGAGAGATGGAACACACAGCGTGTGAGAAATGCACAAAGCCCTGT 1164
Oy 329 ----- 329
Db 1165 GCCCAGTGTGCTATGTGTGGGATGAGCATTGGCAGAGAGTGAAGGCACTTACCAAT 1224
Oy 329 ----- 329
Db 1225 GCCAATATCCAGAGATTTGCTGCTGCTGCAAGAAGATCTTTGGAGGCTGGCAATTTCTGCCG 1284
Oy 329 ----- 329
Db 1285 GAGAGCTTTGATGGGACCCAGACCTCCAAACTGCCCGCTCCAGACAGAGAGCTCCAA 1344
Oy 329 ----- 329
Db 1345 GTGTTGAGACTCTGGAAGAGATCACAGGTTACTATACATCTCAGCATGGCCGAGACAGC 1404
Oy 329 ----- 329
Db 1405 CTGGCTGACCTACGGCTCTTCAGAAACCTGCAGATATCCGGGGAGAGAAATTTCTGCACAT 1464
Oy 329 ----- 329
Db 1465 GCGCCTACTCGTACCTCGAAGGGGTGGGCAATCAGCTGGGTGGGCTGCCCTACATG 1524
Oy 329 ----- 329
Db 1525 AGGGAACGTGGGAGTGGAGCTGGCCCTCATCCACATTAACCCACCTGTGCTTGTGTGAC 1584
Oy 329 ----- 329
Db 1585 ACGGTGCCCTGGGACAGCTTTTGGAAACCGGCAACAGCTGTGCTCCACACTGCCAAC 1644
Oy 329 ----- 329

Db 1645 CGGCCAGAGACAGACTGTGTGGGGCAGAGGCGCTGGCTGCCACACAGCTGTGCCCGGAGG 1704
Oy 329 ----- 329
Db 1705 CACTGTGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCAGATTCCTTGGGGCAG 1764
Oy 329 ----- 329
Db 1765 GAGTCCGTGGAGAAATGCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATCCAGG 1824
Oy 329 ----- 329
Db 1825 CACTGTTGCCGTGCCACCCTGAGTGTGACGCCAGAAATGCTCAGTACCTGTTTGA 1884
Oy 329 ----- 329
Db 1885 CCGGAGGCTACCAAGTGTGTGCGCTGTGCCCACTAATAAGACCTCCCTTGTGCGTGGC 1944
Oy 329 ----- 329
Db 1945 CGCTGCCCGAGCGGTGTGAACCTGACCTCTCTACATGCCCATCTGGAATTTCCAGAT 2004
Oy 329 ----- 329
Db 2005 GAGGAGGCGCATGCCAGCTTGCCCATCAACTGCACCACTCCTGTGTGAGACTGGAT 2064
Oy 329 ----- 329
Db 2065 GACAAAGGCTGCCCGCGAGCAGAGACAGCCAGCCCTGTGAGTCCATGCTCTGCGGTG 2124
Oy 329 ----- 329
Db 2125 GTTGGCATTTCTGCTGTGTGTCTTTGGGGGTGTGTTTGGATCCTCATCAAGGAGCG 2184
Oy 329 ----- 329
Db 2185 CAGCAGAAGATCCGGAAGTACACAGATGCGAGACTGTGCAGAAACGAGGCTGTGGAG 2244
Oy 329 ----- 329
Db 2245 CCGCTGACACCTAGCGGAGCCAGTGCACCAACAGGCGCAGATCCGATCCTTAAAGAGAG 2304
Oy 329 ----- 329
Db 2305 GAGCTGAGGAAGTGAAGTGTGTTGATCTGGCGCTTTGGCAGAGTCTACAAAGGCATC 2364
Oy 329 ----- 329
Db 2365 TGGATCCCTGATGGGAGATGTGAATAATCCAGTGGCCATCAAGTGTGAGGGAAC 2424
Oy 329 ----- 329
Db 2425 ACATCCCCCAAGCCACAAAGAAATCTTAGACGAAGCATACGTATGCTGTGTGGC 2484
Oy 329 ----- 329
Db 2485 TCCCATATGTCTCCCGCTTGTGGCATCTGCTGACATCCAGGGTGCAGTGTGACA 2544
Oy 329 ----- 329
Db 2545 CAGCTTATGCCCTATGCTGCTCTTAGACATGTCCGGAAACCGCGGAGCGCTGGGC 2604
Oy 329 ----- 329
Db 2605 TCCAGAGACTGCTGAACCTGGTGTATGCAGATTGCCAAGGGAGTACGTACCTGGAGAT 2664
Oy 329 ----- 329
Db 2665 GTGCGGTGTACACAGGAGCTTGGCGGCTGGAACGTGTGTCAAGAGTCCAAACAT 2724
Oy 329 ----- 329
Db 2725 GTCAAAATTAACAGACTTGGGCTGGCTGGCTGTGACATTGACAGACAGATGACAT 2784

```

OY 329 ----- 329
Db 2785 GCAGATGGGGCAAGTGGCCCATCATAGTGGAGCGCTGGAGTCCATTCTCCGGCGG 2844
OY 329 ----- 329
Db 2845 TTCACCCACCAAGAGTGTGTGTGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTT 2904
OY 329 ----- 329
Db 2905 GGGGCAAACTTACATGATGGATGCCAGCCGGGAGATCCCTGACCTGTGGAAAAGGG 2964
OY 329 ----- 329
Db 2965 GAGCGGCTGCCCGCCCGCCCATCTGCACATGATGTCTACATGATGATGTCATAATGT 3024
OY 329 ----- 329
Db 3025 TGTGATGATTGACTCTGAATGTGCGGCAGATTCGGGAGTTGTGTCTGAATTCTCCGC 3084
OY 329 ----- 329
Db 3085 ATGGCCAGGAGACCCCGCCGCTTGTGTGTATCCAGATGAGACTTGGCCAGCCAGT 3144
OY 329 ----- 329
Db 3145 CCCTTGACAGACACTTCTACCCGCTCAGCTGTGAGAGACGATGACATGGGGACCTGTG 3204
OY 330 ----- 330
Db 3305 GATCCTGAGAGATTTCTGTGTAACCCAGAGGGCTTCTTCTGTCCAGACCTCCCGCGG 3264
OY 331 AlAglYlmeTValhShlSarghlsArgSerSerThrArgSerGlyGlyAsp 330
Db 3265 GCTGGGGGCAATGTCACACAGAGGACGGAGCTCATCTACAGAGATGGGGGGGAG 3324
OY 351 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370
Db 3325 CTGACACTAGAGGCTGAGGCGCTTGAGAGAGAGGCGCCAGGTCTCCACTGCGCCCTCC 3384
OY 371 GluGlyAlaGlySerAspValAlaPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeu 390
Db 3385 GAAAGGGGCTGCTCCGATGATTTGATGCTGACCTGGGAATGGGGGCAAGGAGGCTG 3444
OY 391 GluSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal 410
Db 3445 CAAAGCCTCCCGACACATGAGCCCGCTCTACAGGGGTACAGTGAAGAGCCACAGTA 3504
OY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
Db 3505 CCCCCTGCTCTGAGACTGATGCTGACCTGACCTGACCTGACCTGACCTGACCTGAA 3564
OY 431 TyrValAsnGlnProAspValAlaArgProGlnProProSerProArgGluGlyProLeuPro 450
Db 3565 TATCTGAACCAAGCAGATGTTGGCGCCCGCCAGCCCTTCCCGCAGAGGGCCCTCTGCT 3624
OY 451 AlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsn 470
Db 3625 GCTCCCGACCTGCTGTGCTGCTGCTGAAAGGCGCAAGACTCTCTCCCGAGGAGACAT 3684
OY 471 GlyValValLysAspValAlaPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
Db 3685 GGGGTCTCTCAAGAGCTTTTGGCTTTGGGGGTGCTGCGGAAAGCCCGGATCTTACA 3744
OY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
Db 3745 CCCAGGAGAGAGTGGCTCCCTGACGCCACCTCTCTCTGCTTCAAGCCGCTTGGAC 3804
OY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
Db 3805 AACCTTATATCTGAGACCAAGAGCCACAGAGAGGGGGGTCCACACAGAGACTTCAAA 3864

```

```

OY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaPro 550
Db 3865 GGGACACCTACGGCAGAAACCCAGAGTACTGCTGTGACGTCACA-----GTGTGA 3918
OY 551 AlaArgSerProSerPro 556
Db 3919 ACCAGAAAGCCAGTCCG 3936

RESULT 7
LOCUS BD005474 4530 bp DNA linear PAT 31-JAN-2002
DEFINITION Cellular immunogens useful cancer vaccines.
ACCESSION BD005474
VERSION BD005474.1 GI:18633845
KEYWORDS JP 2001501909-A/4.
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE
1 (bases 1 to 4530)
AUTHORS Halpern, M.S. and England, J.M.
TITLE Cellular immunogens useful cancer vaccines
JOURNAL Patent: JP 2001501909-A 4 13-FEB-2001;
ALLEGHENY UNIVERSITY OF THE HEALTH SCIENCES

COMMENT
OS Unidentified
PN JP 2001501909-A/4
PD 13-FEB-2001
PF 13-JAN-1997 JP 1997526124
PR 19-JAN-1996 US 60/010262
PI MICHAEL S HALPERN, JAMES M ENGLAND
PC A01K63/00, A61K39/00, A61K39/38, A61K48/00, C12N5/00, C12N15/00 CC

Strandedness: Single;
Topology: Linear;
FH Key 1.4530 Location/Qualifiers
FT source 1.4530 /organism='unidentified'.

FEATURES
source 1.4530 Location/Qualifiers
BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN
Alignment Scores: 1.49e-80 Length: 4530
Pred. No.: 2423.00 Matches: 524
Score: 41.14% Conservative: 21
Best Local Similarity: 40.75% Mismatches: 736
Query Match: 64.80% Indels: 4
Gaps: 4
DB: 6

US-09-821-883-2 (1-690) x BD005474 (1-4530)
OY 2 ArgAlaAlaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 94 CGGCCCTCCACCGGGGTCCAGCGGAGCCATGCGCGCGAGCGAGCGAGCAGCATG 153
OY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGlyLeuAlaArg 37
Db 154 GAGCTGGCGGCTGTGCTGCCCTG-----GGGCTCTCTCTGCGCTTGTGCCCCC 204
OY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 205 GGAGCCCGCAGACCCCAAGTGTGTCACCGGCAACATGAAAGTGGCGCTGCTGCGCAGT 264
OY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValAlaGlnGly 77
Db 265 CCCAGAGACCACTGAGACATGCTCCGCCACTCTACACAGGGCTCCAGTGGTGCAGGA 324
OY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGGAAGTCACTACCTGCGCCACCAATGCCAGCTGCTTCTGCGAGGATATCCAG 384
OY 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117

```

Db 385 GAGGTGCAGGGCTACGTCTACGCTCACAACCAAGTAGAGCAGGTCCCACTGCAGAG 444
 Oy 118 LeuAlaIleValArgIleThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
 Db 445 CTGGCGATTGTGCGAGGACCCAGCTCTTTGAGGACAACTAAGCCCTGGCTGTAGAC 504
 Oy 138 AsnGluAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
 Db 505 AATGAGAACCCGCTACACATACACCCCTGTGCACAGGGGCTCCCGCAGAGGCTGGCG 564
 Oy 158 GluLeuGlnLeuArgSerLeuThrGlnIleLeuLysGlyValLeuIleGlnArgAsn 177
 Db 565 GAGCTGCAGCTTCGAAGCCTCACAGAGATCTTGAAGAGAGGGGCTCTGTGATCCAGCGAAC 624
 Oy 178 ProGlnLeuGlySerGlyAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln 197
 Db 635 CCGCCAGCTCTGTACACAGACACGATTTGTGGAAGGACATCTTCCACAGAACCAACAG 684
 Oy 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
 Db 685 CTGGCTTCACACATATAGACACCAACGCTCTGGGCTGCACCCCTGTCTCCGANTG 744
 Oy 218 CysLysGlySerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThr 237
 Db 745 TGTAAAGGCTCCCGCTGCTGGGGAGAGATCTGAGAGATTTCTCAGAGCCTGACGCGCACT 804
 Oy 238 ValCysAlaGlyIleGlyCysAlaArgCysLysGlyProLeuProThrAspCysGlyHisGlu 257
 Db 805 GCTCTGCGGGGTGGCTGTGGCCCGCTGCAGAGGGGCCACTGCCACTGACGTGCTCCATGAG 864
 Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
 Db 865 CAGTGTGGTGGCGGGGTGCAGCGGGCCCAAGCACTGTGATGCTGTGGCTGTCCACTTC 924
 Oy 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
 Db 935 AACCCAGAGGCGATCTGTAGCTGCACCTGCCAGCCCTGTCTACACACAGACAGACAG 984
 Oy 298 PheGluSerMetProAsnProGluArgTyrThrPheGlnAlaSerCysValThrAla 317
 Db 965 TTTGAGTCATGCCCAATGCCAGGGCCGGATATACATTTGGCGCCAGCTGTGTGACTGCC 1044
 Oy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer 329
 Db 1045 TGTCCCTACAACTACTCTTCTACGAGCGTGGATCTGCACCCCTGTGCCCCCTGCAC 1104
 Oy 329 329 329
 Db 1105 AACCAAGAGGTGACAGACAGAGATGAAACACAGCGGTGTGAGAGTGCACAAAGCCCTGT 1164
 Oy 329 329 329
 Db 1165 GCCCAGTGTGCTATGTGCTGGGCATGGACACTTGGCAGAGGTGAGGGCAGTTACCACT 1224
 Oy 329 329 329
 Db 1225 GCCAATATCCAGAGATTGTGCTGGCTGCAGAAAGATCTTTGGAGGCTGTGCTGCGG 1284
 Oy 329 329 329
 Db 1285 GAGAGCTTTGATGGGACCCAGCTCCAAACACTGCCCGCTCCAGCCAGAGCAGCTCCA 1344
 Oy 329 329 329
 Db 1345 GTGTTTGAAGACTCTGTGAAGAGATCACAGGTTACTATACATCTCAGCATGGCCGAGACAC 1404
 Oy 329 329 329
 Db 1405 CTGGCTGACCTCAGGCTTCTCCAGAACTGCAGAAATATCCGGGAGAGAAATTCTGCACAAT 1464
 Oy 329 329 329

Db 1465 GCGCCCTACTGCTGATCCCTGCAAGGCTGGGCATACAGCTGGCTGGGCTCACTG 1524
 Oy 329 329 329
 Db 1525 AAGGAAGTGGGAGTGGAGCTGGCCCTCATCCACATTAACCCACCTCTGCTTGTGCAC 1584
 Oy 329 329 329
 Db 1585 ACCGTGCCCTGGGACACAGCTTTTGGAAACCGCACCAAGCTCTGCTCACACATGCCAAC 1644
 Oy 329 329 329
 Db 1645 CGGCCAGAGACGAGTGTGTGGCGAGGGGCTGGCCCTGCACACAGCTGTGCGCCGAGGG 1704
 Oy 329 329 329
 Db 1705 CACTGCTGGGGTCCAGGGCCACCCAGTGTGTCAACTGCAGCAGTCTCTTCGGGGCCAG 1764
 Oy 329 329 329
 Db 1765 GAGTGCATGAGAGAAATGCCAGTACTGCAGAGGGGCTCCCGAGGAGTATGTGAATGCAGG 1824
 Oy 329 329 329
 Db 1825 CACTGTTGCCCTGCCACCCCTGAGTGTGAGCCCGAGAAATGGCTCAGTACACTGTTTGA 1884
 Oy 329 329 329
 Db 1885 CCGGAGGCTGACCAAGTGTGTGGCTGTGGCCCACTATTAAGAACCTTCCTTGTGCTGGCC 1944
 Oy 329 329 329
 Db 1945 CGCTGCCCGAGGCGTGTGAACCTGACCTCTCTACATGCCATCGGAATTTCCAGAT 2004
 Oy 329 329 329
 Db 2005 GAGGAGGGCGCATGCCAGCCTTGCCCATCAACTGACCCACTCTGTGTGAGACTGTGAT 2064
 Oy 329 329 329
 Db 2065 GACAAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTGTGACCTCATGCTCTGTGGGTG 2124
 Oy 329 329 329
 Db 2125 GTTGCATTTCTGTGCTGTGTGCTTTGGGGGTGTGCTTTGGGATCTCATCAAGCAGCG 2184
 Oy 329 329 329
 Db 2185 CAGCAGAAAGATCCGGAAGTACACGATGCGGAGACTGTGCAGAAACGGAGCTGTGGAG 2244
 Oy 329 329 329
 Db 2245 CCGCTGACACCTAAGCGGAGCGATGCCCAACAGGGCGAGATCGGATCCTGAAAGAGAG 2304
 Oy 329 329 329
 Db 2305 GAGCTGAGAAAGTGAAGGTGCTTGATCTGCGCTTTTGGACAGTCTACAAGAGCATC 2364
 Oy 329 329 329
 Db 2365 TGGATCCTGATGAGGGAGAAATGTGAATAATTCAGTGGCCATCAAGTGTGAGGAGAAAC 2424
 Oy 329 329 329
 Db 2425 ACATCCCCCAAGCCAAACAGAAATCTTAGACGAAGCATACGTATGCTGTGTGGGC 2484
 Oy 329 329 329
 Db 2485 TTCCCATATGTCTCCGCTTGTGGCATGTGCTTGACATCCAGGTTGCAATGGTGTGACA 2544
 Oy 329 329 329
 Db 2545 CAGCTTATGCCCTATGAGTGTGCTTGTAGACATGTGCCGGAACCGGAGCGCTGGGC 2604

```

QY 329 ----- 329
Db 2605 TCCGAGGACCTGCTGAAGTGTATGACATTTGCCAAGGGGATGAGTACTGAGAGAT 2604
QY 329 ----- 329
Db 2665 GTGGGCGCTGACACAGGAGACTTGGCGGCTGGAAAGCTGTGTCAAGAGTCCCAACCAT 2724
QY 329 ----- 329
Db 2725 GTCAAAATTACAGACTTCGGGCTGGCTGGCTGACATTGACAGACAGAGTACCAT 2784
QY 329 ----- 329
Db 2785 GCAGATGGGGGCAAGTGCCCATCAAGTGTGAGGCTGGAGTCCATTCTCCGGCGG 2844
QY 329 ----- 329
Db 2845 TTACACCCACAGAGTGTGTGTGAGATTATGTGTGACTGTGTGGAGTGTACTTTT 2904
QY 329 ----- 329
Db 2905 GGGGCAAACTTACGATGGATGCCAGCCGGGAGATCCCTGACTGCTGGAAGGGG 2964
QY 329 ----- 329
Db 2965 GAGGCGTCCCCAGCCCCCATCTGACATTTGATGTATCATGATCATGTCAAAATGT 3024
QY 329 ----- 329
Db 3025 TGATGTATGACTGTGAATGTGCGGCAAGATTCCGGAGTGTGTGAATTTCTCCGC 3084
QY 329 ----- 329
Db 3085 ATGGCCAGGAGCCCAAGCGCTTTGTGTGATCCAGAAATGAGACTTGGGCCAGCCACT 3144
QY 329 ----- 329
Db 3145 CCCTTGGACAGACCTTTCTACCGCTCACTGCTGGAGAGCATGATGGGGAGCTGTG 3204
QY 330 ----- Gly 330
Db 3205 GATGCTGAGGAGTATGTGGTATCCACAGAGGCTTCTTGTCTCAGACCTGCCCGGCG 3264
QY 331 AlaGlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyAsp 350
Db 3265 GCTGGGGGCAATGCTCCACACAGGACCGGACGCTCATCTACACAGAGTGGCGGTGGGAC 3324
QY 351 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 370
Db 3325 CTGACACTAGGGCTGAGAGCCCTCTGAAGAGAGGAGCCGCCAGGCTCCACTGGCACCTCC 3384
QY 371 GluIleValIleGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeu 390
Db 3385 GAAAGGGGCTGGCTCGAATGATTTGATGTGACTGGGAATGGGGAGGCAAGGGGCTG 3444
QY 391 GluSerLeuProThrHisAspProSerProLeuGluAlaArgTyrSerGluAspProThrVal 410
Db 3445 CAAAGCCTCCACACATGAGACCCAGCCCTTACAGCGGTACAGTACAGAGAGCCACAGTA 3504
QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
Db 3505 CCCCTGCCCTCTGAGACTGATGCTAGCTGAGTCCGCCCTGAGACTGACAGCCGCCAGCTGAA 3564
QY 431 TyrValAsnGlnProAspValArgProGluProProSerProAlaGlyGluGlyProLeuPro 450
Db 3565 TATGTGAACACAGCCAGATGTTGGGCGCCAGCCCTTGGCCCGAGAGGGGCGCTCTGCTCT 3624
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaArgAlaIleThrLeuSerProGlyIleAsn 470
Db 3625 GCTGCCGAGACTGCTGTGGTGCACACTCTGGAAGAGGCCAAGACTCTCTCCCAAGGAGAAAT 3684

```

```

QY 471 GlyValIleValIleAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
Db 3685 GGGGTGCTGAAGACGTTTTCCTTTGGGGGTCCGCTGGAGAACCCGAGTACTTGACA 3744
QY 491 ProGlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAsp 510
Db 3745 CCCCAGGAGAGAGCTGCCCTCCACGCCCTCTCTGCTTGCACGCCAGCTTGGAC 3804
QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
Db 3805 AACCTCTATTACTGGACACAGAGACCCACAGAGGGGGGCTCCACCCAGCACCTTCAAA 3864
QY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro 550
Db 3865 GGGACACTACGGCAGAGAACCCAGAGTACTGGTGTGAGACGTGCCA-----GTGTGA 3918
QY 551 AlaArgSerProSerPro 556
Db 3919 ACCAGAGGCCCAAGTCCG 3936

RESULT 8
LOCUS 121124 4530 bp DNA linear PAT 07-Oct-1996
DEFINITION Sequence 9 from patent US 551885.
ACCESSION 121124
VERSION 121124.1 GI:1601478
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4530)
AUTHORS Razindin and Sarkar, F.H.
TITLE ERB2 promoter binding protein in neoplastic disease,
JOURNAL Patent: US 551885-A 9 21-MAY-1996;
FEATURES
source 1..4530
location/Qualifiers
BASE COUNT 922 a 1382 c 1346 g 880 t
ORIGIN
Alignment Scores:
Pred. No.: 1,49e-80 Length: 4530
Score: 2423.00 Matches: 524
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 21
Query Match: 64.80% Indels: 736
Gaps: 4
US-09-821-883-2 (1-690) x (121124 (1-4530))
QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 94 CCGGCTCCACAGCGGGGTCCAGCCGAGCATGGGCGCGAGCGGAGAGACCATG 153
QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
Db 154 GAGCTGGCGGCTGTCCGCTGG-----GGGCTCCCTCCGCCCTCTTCCGCC 204
QY 38 GlyValAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 205 GAGCGCGGAGACCCCAAGAGTGTGCACCGGACACACATGAAAGCTGGCGCTCCCTCCAGT 264
QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValIleGlnGly 77
Db 265 CCGGAGAGCCACCTGAGACATGCTCGGCACTCTACAGAGGCTCCAGGTGTGACAGGA 324
QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGGAACCTCACTACCTGCGCCACCAATGCCAGCTGTCTTCCGTCAGAGATATCAG 384
QY 98 GluValGlnGlyTyrValIleValIleAlaHisAsnGlnValIleArgGlnValIleProLeuGlnArg 117
Db 385 GAGGTGAGGGGCTACGTGCTCATCTGTCACAAACCAAGTGAAGGAGGTGCCACATGCAAGAG 444

```

OY	118	Leu1Arg11LeuValArgGlyThrThrGlnLeuPheGlnIAspAsnTyrAlaLeuAlaValLeuAsp	137
Db	445	CTGGGAGATTGGCCAGAGCCACCAGCTCTTTGAGAGCAACTATGAGCCCTGGCCGTGTATAC	504
OY	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
Db	505	AATGAGACCCGGCTGAAACATATACCACCCCTGTGCAGAGGGGCTCCCCAGAGAGCCCTGGG	564
OY	158	GluLeuGlnLeuArgSerLeuThrGluIleLeuValGlyGlyValLeuIleGlnArgAsn	177
Db	565	GAGGTGACGCTTCACACCCCTCACAGAGATCTTGAAAGAGGGGCTGTGATCCAGGGAAAC	624
OY	178	ProGlnLeuGlyTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln	197
Db	625	CCCCAGCTCTGCTACAGAGACAGATTTTGTGAAAGACATCTTCCACAAAGAACACAG	684
OY	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet	217
Db	685	CTGGCTCTCACACTGATAGACACCAACCGGCTCTGGGGCTGGCACCCCTGTTCTCCGATG	744
OY	218	CysLysGlySerArgCysTyrPrgLysIleuSerSerGluAsCysGlnSerLeuThrArgThr	237
Db	745	TGTAAAGGCTCTCCGCTCTGGGAGAGAGTTCTGAAGATTGTACAGGCTGACGGGCACT	804
OY	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu	257
Db	805	GTCGTGGCCGGTGGCTGTGGCCGCTGGAGAGGGGCCACTGCCCACTGACTGTCGCCATGAG	864
OY	258	GlnCysAlaAlaArgLysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277
Db	865	CAGTGTGCTGGCGGCTCTACGGGGCCCCAAACACTGTGACTGGCTGGCCCTCCACTTC	924
OY	278	AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297
Db	925	AACCAACAGTGGCACTGTGAGCTGTGCATGGCCAGCCCTGTCACTACAAACACAGACAG	984
OY	298	PheGlnSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla	317
Db	985	TTTGTAGTCCAAAGCCCAATCCCGAGGGCCGGTAAATACATTTGGGGCCAGCTGTGACTGCC	1044
OY	318	CysProTyrAsnTyrLeuSerThrAspValGlySer-----	329
Db	1045	TGTCTCTACACTACCTTTCTACGGAGCTGGATCTTGCAACCTGTGTCGCCCTCTGAC	1104
OY	329	-----	329
Db	1105	AACCAAGAGTGACAGACAGAGAGATGGAACACAGCGGTGTGAGAAAGTCAGCAAGCCGTGT	1164
OY	329	-----	329
Db	1165	GCCCGAGTGTGCTATGTGCTTGGGCATGAGACACTTGCAGAGAGTGAAGGCAAGTTACAGT	1224
OY	329	-----	329
Db	1225	GCCAAATATCCAGAGAGTTTGGCTGGCTGCAAGAAATCTTTGGAGAGCTTGGCATTTCTGGCG	1284
OY	329	-----	329
Db	1285	GAGAGCTTTGATGGGAGCCAGCGCTCCAAACACTGCCCCCTCCAGCCAGAGAGAGCTCCAA	1344
OY	329	-----	329
Db	1345	GTTGTTTGAAGACTCTGGAAGAGATCACAGGTTACCTTATCATCTCAGCATGGCCGGACAGC	1404
OY	329	-----	329
Db	1405	CTGCCTGACTCAGACGTCCTTCCAGAACTCAGAGTATATCCGGGAGAGATTTCTGCACAAAT	1464
OY	329	-----	329
Db	1465	GAGCGCTACTGCTGACCCCTGCAGAGGGCTGGGAGTACGTGCTGGGGCTGCGCTCACTG	1524

0Y	329	-----	329	-----	329
Db	1525	AGGGAACGGGACAGTGGATGGACCGCCCTCATTCACACATAAACCCACCTCTGCTGTGCAC	1584		
0Y	329	-----	329	-----	329
Db	1585	ACGGTGCCTTGGGACAGCTCTTTGGGAACCGGACCAAGCTCTGCTCCACACGTCCAAAC	1644		
0Y	329	-----	329	-----	329
Db	1645	CGGCCAGAGACAGATGTGTGGGCGAGGGCCTGGCCCTCCACACAGCTGTGGCCGAGGG	1704		
0Y	329	-----	329	-----	329
Db	1705	CACGTCTGGGGTCCAGGGCCACCACAGTGTTCACACTGCAGCAAGTTCTTGGGGCCAG	1764		
0Y	329	-----	329	-----	329
Db	1765	GAGTCCGTGGAGAAATCCAGATACATCCAGGGGCTCCCGAGGAGTATGTGAATGCCAGG	1824		
0Y	329	-----	329	-----	329
Db	1825	CACGTTTGCCGCTGCACACCTGAGTGTCAAGCCCCAGAAATGGCTCAGTGCCTTTTGA	1884		
0Y	329	-----	329	-----	329
Db	1885	CCGGAGGCTGACAGTGTGTGGCCTGTGCCACTATAGAGACCTCCCTTTCCTGGCC	1944		
0Y	329	-----	329	-----	329
Db	1945	CGCTGCCCAAGCGGTGTGAACCTGACCTCTCCTACATGCCCATCTGGAATTTCCAGAT	2004		
0Y	329	-----	329	-----	329
Db	2005	GAGGAGGGCGCATGCGACGCTTGCCCATCAACTGCACCCACTCCGTGTGTGACCTTGAT	2064		
0Y	329	-----	329	-----	329
Db	2065	GACAAGGCGTCCCGCCGAGAGAGAGCCAGCCCTGTGACGTGCATGTCTTCCGGTG	2124		
0Y	329	-----	329	-----	329
Db	2125	GTTGGCATTTCTGTGTGTGTCTTGGGGGTGTCTTTGGGATCCTCATCAAGCAGCG	2184		
0Y	329	-----	329	-----	329
Db	2185	CAGCAGAAGATCCGGAAGTACAGATGCGGAGACTCTCTCAGAGAAACGAGCTGTGGAG	2244		
0Y	329	-----	329	-----	329
Db	2245	CCGCTGACACCTAGCGAGCGATGCCCAACAGAGCGCAGATGCGATCCTGAAGAGACG	2304		
0Y	329	-----	329	-----	329
Db	2305	GAGCTGAGAGAGTGGAAGGTCTTGATCTGGCGCTTTTGGCAGAGTCTACAAGGGCATTC	2364		
0Y	329	-----	329	-----	329
Db	2365	TGGATCCCTGATGGGGAGAAATGTGAATAATTCAGTGGCCATCAAAAGTGTGAAGGAAC	2424		
0Y	329	-----	329	-----	329
Db	2425	ACATCCCCCAAGCCACAAGAAATCTTAGAGGAAGCATACGATGGCTGTGGGC	2484		
0Y	329	-----	329	-----	329
Db	2485	TCCCATATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCAGGGTGACCTGTGACA	2544		
0Y	329	-----	329	-----	329
Db	2545	CAGCTATGCCCTATGGCTGCCCTTAGACATGTCCGGAAAAACCGGAGCGCTGTGGC	2604		
0Y	329	-----	329	-----	329

Db 2605 TCCAGAGACCTGCTGAAGTGTGTATGACAGATTGCCAAGGGAGATGACCTACTGAGAGAT 2664
 QY 329 ----- 329
 Db 2665 GTGGGCTCGTACACAGAGAGACTTGGCGCTCGGAAGCTGCTGTCAGAGAGTCCCAACCAT 2724
 QY 329 ----- 329
 Db 2725 GTCMAAATTACAGACTTCGGGCTGGCTCGGCTGTGACATTCAGACAGAGACTTACCAT 2784
 QY 329 ----- 329
 Db 2785 GCAGATGGGGCAAGTGGCCCATCAAGTGGATGGCTGTGAGTCCATTCTCCGCGCGG 2844
 QY 329 ----- 329
 Db 2845 TTCACCCACAGAGTATGTGTGTGAGATTATGTGTGACTGTGTGGAGAGTATGACTTTT 2904
 QY 329 ----- 329
 Db 2905 GGGGCCAAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACTGTGAAAGGGG 2964
 QY 329 ----- 329
 Db 2965 GAGCGGCTGCCCCAGCCCCCATCTGCACATTGATGTATCATGATGATGTCATAATGT 3024
 QY 329 ----- 329
 Db 3025 TGGATGATGACTCTGAATGTGGCCCAAGATTCCGGGAGTTGGTGTGAATTCTCCGC 3084
 QY 329 ----- 329
 Db 3085 ATGGCCAGGAGACCCAGCGCTTGTGTGATCATCAGATGAGACTTGGCGCCAGCCAGT 3144
 QY 329 ----- 329
 Db 3145 CCCTTGACAGACACTCTACCGCTCAGCTGTGAGAGACATGACATGGGGAGCTGTG 3204
 QY 330 ----- Gly 330
 Db 3205 GATGCTGAGAGATATCTGTATCCACAGAGAGGCTTTCTGTCCAGACCTGCCCGGGC 3264
 QY 331 AAGGlylymetvalhshisatghisatgsersestheratgserglylyasp 350
 Db 3265 GCTGGGGCATGTGCCACAGAGACCGCATCTACAGAGTGGCGGGGAGC 3324
 QY 351 LeuThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSer 370
 Db 3325 CTGACACTAGGGCTGGAGCCCTCTGAAGAGAGAGCCGCCAGGTCTCCACTGGCACCTCC 3384
 QY 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaGlyLeu 390
 Db 3385 GAAGGGGCTGGCTCCGATGATTTGATGTGACTGGGAATGGGGAGCAGCAGGGGCTG 3444
 QY 391 GluSerLeuProThrHisAspProSerProLeuGluAlaArgTyrSerGluAspProThrVal 410
 Db 3445 CAAGGCTTCCACACATGAGCCCAAGCCCTCTACAGGGGAGACAGTGGAGCCACAGTA 3504
 QY 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
 Db 3505 CCCCTGCCCTGAGTATGAGTGGTACGTGGTGGCCCTGACCTGACCTGACCCCGAGAA 3564
 QY 431 TyrValAsnGluProAspValArgProGluProProSerProArgGluGlyProLeuPro 450
 Db 3565 TATGTAAACAGCAGATGTTGGCCAGCCCTCGCCGAGAGAGGGCCCTGCTCCT 3624
 QY 451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaAlaGlyThrLeuSerProGlyLysAsn 470
 Db 3625 GCTGCCGAGCTCTGTGTGACTGTGAAAGGGCCCAAGACTCTCTCCAGGAGAAAT 3684
 QY 471 GlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThr 490
 |||||||

Db 3685 GGGGCTGCTCAAGACGTTTTTGCCTTTGGGGGTGCCGTGGAGAACCCGAGTACTTGACA 3744
 QY 491 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
 Db 3745 CCCAGGAGAGAGTGGCCCTCAGACCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3804
 QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys 530
 Db 3805 AACCTTATTACTGAGGACAGGAGACCCACAGAGCGGGGCTCCACCCAGACTTCAAA 3864
 QY 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaPro 550
 Db 3865 GGGACACTTACGGCAGGAACCCAGAGTACCTGGCTGTGACGTGCCA-----GTGTGA 3918
 QY 551 AlaArgSerProSerPro 556
 Db 3919 ACCAGAGGCCAAGTCCG 3936
 RESULT 9
 159745
 LOCUS 159745 4530 bp DNA linear PAT 07-OCT-1997
 DEFINITION Sequence 9 from patent US 5654406.
 ACCESSION 159745
 VERSION 159745.1 GI:2478377
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 4530)
 AUTHORS Razuddin and Sarkar, F. Hogue.
 TITLE Antibody to ERBB2 promoter binding factor
 JOURNAL Patent: US 5654406-A 9 05-AUG-1997;
 FEATURES
 source 1..4530
 BASE COUNT 922 a 1382 c 1346 g 880 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,49e-80 Length: 4530
 Score: 2423.00 Matches: 524
 Percent Similarity: 41.14% Conservative: 5
 Best Local Similarity: 40.75% Mismatches: 21
 Query Match: 64,808 Indels: 736
 DB: Gaps: 4
 US-09-821-883-2 (1-690) x 159745 (1-4530)
 QY 2 ArgAlaAlaProLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17.
 Db 94 CGGCGCTCCAGCGGGGTCCAGCGGAGCATGGGCGCGAGCGGAGCGAGTGAACCATG 153
 QY 18 PheLeuPheLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37.
 Db 154 GAGCTGGCGGCTTGTGCCCTG-----GGCTCTCTCTCGGCCCTCTTGGCCCCC 204
 QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57.
 Db 205 GGAGCCCGAGCACCCCAAGGTGACCGGACAGACATGAAGTGGCGCTCCCTGCCAGT 264
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77.
 Db 265 CCCGAGACCCACCTGGACATAGTCCCGCACCTTACACAGGGCTGCGAGGTGCGAGGA 324
 QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97.
 Db 325 AACCTGGAATCACTACCTGCGCCACCAATGCCAGCTGTCTCTCTGCGAGATATCCAG 384
 QY 98 GluValGlnGlyTyrValLeuAlaAlaHisAsnGlnValArgGlnValProLeuGlnArg 117.
 Db 385 GAGGTGAGGGCTACGTCATCGCTCACAACAGTGAAGGAGGAGGTCACCATGCGAGAG 444
 QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
 |||||||

|||||
Db 445 CTGCGATGTTGCGAGGACCCAGCTTTTGAGGACAACTGACCTGGCCGTGCTAGAC 504
Oy 138 AsnGlyAspProLeuAsnAsnThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505 AATGAGAACCCGCTAACAAATACCAACCCCTGTACAGGGGGCTCCCAAGAGCCCTGGCG 564
Oy 158 GluLeuGlnLeuArgSerLeuThrGluLeuLysGlyGlyValLeuIleGlnArgAsn 177
Db 565 GAGCTGACGCTTCGAAAGCTCCACAGATCTTGAAAGAGGGGTCTTATCCAGGGAAAC 624
Oy 178 ProGlnLeuCySTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln 197
Db 625 CCCCAGCTCTGCTACACAGACACGATTTTGTGGAGGACATCTTCACAAAGAAACACAG 684
Oy 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
Db 665 CTGGCTTCACACTGATAGACACCAACCGCTCTCGGCCCTGCCACCCCTGTCTCGATG 744
Oy 218 CysLysGlySerArgCysTrpGlyLysSerGlnAspCysGlnSerLeuThrArgThr 237
Db 745 TGTAAAGGCTCCCGCTGCTGGGGAAGAGTTGTGAGATTTGTCAAGCCTGACGCGCACT 804
Oy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu 257
Db 805 GTCTGTGCGGCTGTGCTGCCGTGCAAGGGGCCACTGCCACTGCTGCTGCATGAG 864
Oy 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865 CAGTGTGTGCGCGGTGACAGGGGCCCAAGCACTGTGACTGCTGCTGCTGCACATTC 924
Oy 278 AsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
Db 925 AACCACTAGGATGTGTGAGTGCATGCGCCAGCTGTGTACTACACACACACACAG 984
Oy 298 PheGlnSerMetProAsnProGlnGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
Db 985 TTTGAGTCCATGCCCCAATCCCGAGGGCCGGTATACATTCGCGCGCAGGTGTGATGCC 1044
Oy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer 329
Db 1045 TGTCCCTACAACTACCTTTCTAGGAGACGTGGATCTGCACCCCTGTGCCCTGCAC 1104
Oy 329 329
Db 1105 AACCAAGAGGTGACAGAGAGATGGAACACAGCGGTGAGAGATGCCAGAACCCCTGT 1164
Oy 329 329
Db 1165 GCCCGAGTGTCTATGTCTGGGCATGAGACATTGCGAGAGTGAGGGCACTTACCACT 1224
Oy 329 329
Db 1225 GCCAATATCCAGAGATTGTGCTGCTGCAAGAAGATCTTTGGAGCGCTGGCATTTCTGCCG 1284
Oy 329 329
Db 1285 GAGACCTTTGATGGGAGACCCAGCCCTCAAACTGCCCCGCTCCAGCCAGAGAGCTCCAA 1344
Oy 329 329
Db 1345 GTGTTTGAGACTCTGAGAGATCAGAGTTACTATACATCTCAGCATGCGCGGAGACG 1404
Oy 329 329
Db 1405 CTGCTGACCTCAGCGCTTTCAGAACTGCAAGTAATCCGGGGAGCAATTTCTGCACAA 1464
Oy 329 329
Db 1465 GGGCGCTACTGCTGACCTGCAAGGGGCTGGGCATCAGTGGCTGGGGCTGGCGTCACTG 1524
Oy 329 329

Db 1525 AGGAACTGGGCGAGTGAAGTGAAGTGGCCCTCATCCACATAACACCCACCTGTGCTGAC 1584
Oy 329 329
Db 1585 ACGTGCCCTGGGACACAGCTCTTTTGGAACCCGACACAAAGCTGTCTCAGACTGCCAAC 1644
Oy 329 329
Db 1645 CGGCCAGAGACAGTGTGTGGCGAGGGCCCTGGCCCTGCACACAGCTGTGGCGCCGAGG 1704
Oy 329 329
Db 1705 CACTGTGGGGTCCAGGGCCACCCAGTGTGTCACTCAGCAGCACTTCTCGGGCGAG 1764
Oy 329 329
Db 1765 GAGTCGTGAGAGAAATGCCAGACTGTGAGAGGGGCTCCCAAGAGATGTCAATGCCAG 1824
Oy 329 329
Db 1825 CACTGTTTGGCGTGCACCTGAGTGTGACGCCCAAGATGGCTCAGTACCTGTTTGA 1884
Oy 329 329
Db 1885 CCGAGGCTGACAGTGTGTGCTGTGCCACTATTAAGAACCCCTTCTGTGCTGGCC 1944
Oy 329 329
Db 1945 CGCTGCCCAAGCGGTGTGAACCTGACCTTCTCTACATGCCCACTGTGAAGTTTCCAGAT 2004
Oy 329 329
Db 2005 GAGGAGGGCGCATGCGCAGCTTGCCCATCAACTGCACCCACTCTGTGTGAGACTGGAT 2064
Oy 329 329
Db 2065 GACAAGGCTGCCCGCCGAGAGAGACCAAGCCCTGTGAGCTCATGCTGTGCGGTG 2124
Oy 329 329
Db 2125 GTTGCAATTTCGTGTGTGTGTGCTTTGGGGGTGTGCTTTGGGATTCCTATCAAGGAGCG 2184
Oy 329 329
Db 2185 CAGCAGAAGATCCGGAAGTACAGATGCGAGACTGTGACAGAAACGAGCTGTGGAG 2244
Oy 329 329
Db 2245 CCGCTGACACTAGCGGAGCGATGCCAACCAAGCGCAGATGCCGATCCTGAAGAGACG 2304
Oy 329 329
Db 2305 GAGCTGAGGAAGTGAAGTGTGCTTGATCTGGCGCTTTTGGCACAGTCTACAAAGGCATC 2364
Oy 329 329
Db 2365 TGGATCCCTGATGGGAGATGTGAATAATCCAGTGGCCATCAAGTGTGAGGGAAC 2424
Oy 329 329
Db 2425 ACATCCCCCAAGCCMACAAAGAAATCTTAGACGAAGCATAGTGTGCTGTGGGC 2484
Oy 329 329
Db 2485 TCCCATATGTCTCCCGCTTCTGGGCATCTGCTGACATCCAGGGTGCAGCTGTGTGACA 2544
Oy 329 329
Db 2545 CAGTTATGCCCTATGCTGCTTATAGACATGTCCGGAACCGGAGCGCTGGGC 2604
Oy 329 329
Db 2605 TCCAGAGACCTGCTGAACTGCTGTATGAGATTGCCAAGGGAGTGAAGTACCTGGAGAT 2664

```

Oy 329 ----- 329
Db 2665 GTGGGCTCGTACAGAGGACTTGGCGCTGGAAAGTGTGTCAGAGTCCAAACCAT 2724
Oy 329 ----- 329
Db 2725 GTCAAAATTACAGACTTCGGGCTGGCTGGCTGTGACATTGACAGAGAGTACCAT 2784
Oy 329 ----- 329
Db 2785 GCAGATGGGGCAAGTGGCCATCAAGTGAAGTGGCGCTGGAGTCCATTCTCCGGCGG 2844
Oy 329 ----- 329
Db 2845 TTCACCCACAGAGTGAATGTGTGAGTTAATGTGTGACTGTGTGGAGAGTATGACTTTT 2904
Oy 329 ----- 329
Db 2905 GGGGCCAAACCTTACGATGGAGTCCAGCCCGGAGATCCCTGACCTGTGAAAAAGGGG 2964
Oy 329 ----- 329
Db 2965 GAGGGCTGCCCGCCACCCCATCTGCACCATTTGATGTCTACATGATCATGTCAATGT 3024
Oy 329 ----- 329
Db 3025 TGGATGATTGACTGTGAATGTGCGCAAGATTCCGGGAGTTGTGTGTAATTTCTCCGC 3084
Oy 329 ----- 329
Db 3085 ATGCCAGAGACCCCGCCGCTTTGTGTGTCATCCAGAAATGAGACTTGGCCCGCAGCCAGT 3144
Oy 329 ----- 329
Db 3145 CCCTTGGACAGACACTTCTACCCGCTCACTGTGAGAGAGCATGATGAGGGGAGCTGTG 3204
Oy 330 ----- Gly 330
Db 3305 GATCCTGAGAGATATCTGTGACCCAGAGGCTTCTTCTGTCCAGACCTGCCCCGGGC 3264
Oy 331 AlaGlyMetValHisHisArgSerSerSerThrArgSerGlyGlyAsp 350
Db 3365 GCTGGGGCATGTCCACACAGGCGCGGCTGATCTATCCAGAGAGTGGGGGAGC 3324
Oy 351 LeuThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer 370
Db 3325 CTGACACTAGGGGCTGAGAGCCCTCTGAAGAGAGGAGGCCCGCCAGCTCTCCAGCTGCC 3384
Oy 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeu 390
Db 3385 GAAAGGGCTGCTCCGATGTATTGTGATGTGACCTGGGAATGGGGGAGCAGCAAGGGGCTG 3444
Oy 391 GluSerLeuProThrHisAspProSerProLeuGluAlaArgTrpSerGluAspProThrVal 410
Db 3445 CAAAGCCCTCCACACATGAGCCCAAGCCCTCTACAGGGGTACAGTGAAGAGCCCAACAGTA 3504
Oy 411 ProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGluProGlu 430
Db 3505 CCCCTGGCCTTGAGACTGATGGCTAGCTTGGCCCTGAGCTGACCTGACCCCGACCTGAA 3564
Oy 431 TyrValAsnGluProAspValAlaArgProGluProProSerProArgGluGlyProLeuPro 450
Db 3565 TATGTGAACGACGAGATGTTGGCGCCAGCCCTTCCGCCCAAGAGGGGCTCTGCT 3624
Oy 451 AlaAlaArgProAlaGlyAlaThrLeuGluAlaArgAlaLysThrLeuSerProGlyLysAsn 470
Db 3625 GCTGCCCGACCTGCTGTGGCCACTCTGAAAGGGGCAAGACTCTCTCCCAAGGGAAGT 3684
Oy 471 GlyValValLysAspValPheAlaPheGlyGlyValAlaValGluAsnProGluTyrLeuThr 490
Db 3685 GGGGTCTCTCAAGAGCTTTTGGCTTTGGGGGCTCCCTGGAGAACCCCGAGTACTTGACA 3744

```

```

Oy 491 ProGluGlyAlaAlaProGluProHisProProAlaPheSerProAlaPheAsp 510
Db 3745 CCCGAGGAGAGCTGCCCTCCAGCCCTCTGCTGCTTCCAGCCAGCTTGCAC 3804
Oy 511 AsnLeuTyrTyrTrpAspGluAspProProGluArgGlyAlaProProSerThrPheLys 530
Db 3805 AACCTTATTACTGGACACAGACCCAGAGGGGGGCTCCACCCAGACTTCAAA 3864
Oy 531 GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro 550
Db 3865 GGGACACTACGGCAGAGAACCAGAGTACTGCTGTGACGTCGCA-----GTGTGA 3918
Oy 551 AlaArgSerProSerPro 556
Db 3919 ACCAGAAGGCCAAGTCCG 3936

RESULT 10
LOCUS HUMER2A 4530 bp mRNA linear PRI 18-SEP-1995
DEFINITION Human tyrosine kinase-type receptor (HER2) mRNA, complete cds.
ACCESSION M11730
VERSION M11730.1 GI:183986
KEYWORDS tyrosine kinase.
SOURCE Homo sapiens (clone: lambda-HER2-436) fetal cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4530)
Cousens,L., Yang-Feng,T.L., Liao,Y.-C., Chen,E., Gray,A.,
McGrath,J., Seeburg,P.H., Libermann,T.A., Schlessinger,J.,
Franke,U., Levinson,A. and Ullrich,A.
Tyrosine kinase receptor with extensive homology to EGF receptor
shares chromosomal location with neu oncogene
Science 230 (4730), 1132-1139 (1985)
2999974
JOURNAL MEDLINE 86070181
PUBMED 2999974
REFERENCE 2 (bases 1701 to 1719)
AUTHORS Ullrich,A.
JOURNAL Unpublished
FEATURES
location/Qualifiers
1..4530
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="lambda-HER2-436"
/dev_stage="fetal"
<1..4530
/product="HER2 mRNA"
151..3918
/note="HER2 receptor"
/codon_start=1
/protein_id="AA75493.1"
/db_xref="GI:306840"
/translation="METALAICRMGLLALLLPGAASNOVCTGDMKRLRASPETHID
MIRHLVDCGVGNLEETPLTPNALSIFDIQEOGVYLIAHNOVROVPLORLRIV
RGTOLEEDNTRALAVLNDGNPLANTTTPYKASPGLELDRSTTELKGVILGRPO
LCYQDTILMDIDIRHNNQALALTLIDNRSFACHPSCPMCGSKWESSDCSLRT
VCAGGARGKRLPTDCHEQCAAGCTGPHSCDLACLNHSIGIELHCPLAVTNT
DTEESMPNEGRTYFGASCVTACPYALSTIDVSCSLVPLAHNOETAEEDTCRCKC
SKPCAVCGLEMEHREYAVTSANIOEPAGCKLFTGSLAFPESEFGDPASTAPL
QEPQLOVFTELLEITGLYVTSAMPDSLPLDSYONLOVIRGLIHNCAVSLTGLGI
SWGLSLRLSGSLALIHNNHLCYHYHYVWDULRNPHQALLRANRPEDCVBEG
LACHOLARGHCWPGPTQVCNCSQFLRQGEVEEERVLQGLREYVNAHCLCPHE
COPQNSVTCFGEPAOCVCAHAKDPFCVAPSGVPRDLISYMPILWFPEDEGACQ
PCPINCTSHVDLDDGCPAEORASPLSTSVAVGILLVVLGVGVGLIKRQOKI
RKTMRRLQETELVPLPESGAMPNOAMRILKETELRKVKVLSGAGSTYKGLMI
POENYKTIYAIKILRENTSPRANKETILDEAYVAGVSPVRLIGLCTSYVLQVT
QLMPYKCLDHLVENRGRIGSDLLMKMOIAKMSYLEVRLVHNDLARNLVASP
NHVKTIDFGIARLIDETELYHADGKVPKIMVALSILRRFTHOSDVWSTGVYWE
LMTFGAKPYDGIAPREIPDLLEGERLPPICITIDVYIMVNCMIIDSECRPFEL
VSEFSRMARPPQRFVYIQNEDLGPASPLDSTFRSLIEDMDKDLDADEAYLVPOGF
FCPDPAPGAGGVNHRHSSTSGGDLTLGLEPEEERPRPLAPSGAGSDVDG
DLGMAKAGLQSLPTIHDPSPLQYSDPVPYPLPSPEIDGYVAPITCSPOEYVNOQVPR
PQPSPREGLPARPAGATLERAKTLISPEKNGVADVFAFGAVENPELTPDGGAA

```


FEATURES	Location/Qualifiers
source	1. .9274

BASE COUNT	2152 a	2538 c	2489 g	2095 t
ORIGIN				

Alignment Scores:

Pred. No.:	3,19e-80	927
Score:	2422.00	528
Percent Similarity:	41.12%	Conservative: 9
Best local Similarity:	60.43%	Mismatches: 31
Query Match:	64.78%	Indels: 738
DB:	6	Gaps: 6

US-09-821-883-2 (1-690) x AX060703 (1-9274)

QY	2	ArgAlaIaIaProLeuLeuLeuAlaIaIaIaIaSerLeuSerLeuGlyPheLeuPheLeu	21
Db	1695	CGCGCCGCGCCCTCCACGCGGGGTCCAGCGGAGCCATGAGACTGGCG-----GCC	1745
QY	22	LeuPhePheTrpLeuAspArgSerValLeuAlaIysGluLeuAlaArgGlyAlaAlaSer	41
Db	1746	TTTGCGCGCGTGG-----GGGCTCTCTTCGCGCCCTTCGCCCCCGGAGCCCGAGC	1796
QY	42	ThrGlnAlaIaIaTrpGlyTrpHisProMetCysLeuArgLeuProAlaSerProGluThrHis	61
Db	1797	ACCCAAAGTGTGACCGGCGACAGACATGAAGTGGGGCTCCGTCCGACAGTCCGAGACCCAC	1856
QY	62	LeuAspMetLeuArgHisLeuTyrgInGlyCysGlnValValGlnGlyAsnLeuAluLeu	81
Db	1857	CTGGACATGGTCCGGCCACCTCTACAGAGGGGTGGCAGAGTGTGCAGAGAACTGGAACTC	1916
QY	82	ThrTyrlaLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGly	101
Db	1917	ACCTACTGGCCCAACAATGGCACCGTCTCTCCACAGAAATCCAGAGAGTGGACGGC	1976
QY	102	TyrValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgTyrIleVal	121
Db	1977	TACGTGCTCATCGTCCACAAACAGTGAAGGAGAGTCCCATGCAGAGCTGGCGGATTGTG	2036
QY	122	ArgGlyThrGlnLeuPheGlnAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspPro	141
Db	2037	CGAGGACACCCAGGCTCTTGAGAGCAACTAGCCCTGGCGGTCTGTAGAAATGAGAACCCG	2096
QY	142	LeuAsnAsnThrTrpProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu	161
Db	2097	CTTGAAACAATACCAACCCCTGTGCACAGGGGCTCCCGAGGGCTCGGGAGCTGCACCTT	2156
QY	162	ArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeuGly	181
Db	2157	CGAAGCTTCACAGAGATCTTGAAAGGAGGGGTCTTATCTCAGCGGAGACCCCACTGTGC	2216
QY	182	TyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr	201
Db	2217	TACCAGAGACAGATTTGTGGAGAGCACTTCCACAAAGAACACAGCTGGCTCTCACA	2276
QY	202	LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer	221
Db	2277	CTGTATAGACACCAACCGCTCTCGGGCTGGCACCCCTGTTCCTCCAGTGTGAAGGGCTCC	2336
QY	222	ArgCysTrpGlyIuSerSerGlnAspCysGlnSerLeuThrArgThrAlaCysAlaGly	241
Db	2337	CGCTGCTGGGGAAGAGATTCTTGAGAGATTGCAGAGCTCAAGCGCACTGTCTGTCCGGT	2396
QY	242	GlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnIuIncysAlaIa	261
Db	2397	GGCTGTGCGCCGCTGCAGAGGGGCCACTGCCACTGACTGCTGCCATAGACAGTGTCTGCC	2456
QY	262	GlyAspThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSerGly	281
Db	2457	GGCTGCACGGGGCCCAAGCACTGTGATGGCTGTGGCTGTCCACTTCAACCAACAGTGGC	2516

Oy	282	IIeYvSGIueuHhIScysProAlaIeuValThrTyrAsnThrAspThrPheGluSerMet	301
Db	2517	ATCTGTGAGCTGCACCTCCCAAGCCCTGGTCACCTACACACAGACAGACTTTGATGCTCATG	2576
Oy	302	ProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyrAsn	321
Db	2577	CCCAATCCCGAGGGCCGGTTACATTTCGGGGCAGCTGTGTACATGCGCTTCACAC	2636
Oy	322	TyrLeuSerThrAspValGlySer	329
Db	2637	TACCTTCTACGAGCGTGGGATCTCCAGCCCTGCTGCCCCCTGCACAAACCAAGAGTG	2696
Oy	329	-----	329
Db	2697	ACAGCAGAGGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGTGCCAGATGCC	2756
Oy	329	-----	329
Db	2757	TATGGTCTGGGACATGAGACACTTGCAGAGGTGAGGGCAGTTACCAAGTGCATATTCAG	2816
Oy	329	-----	329
Db	2817	GAGTTTGTGCTGCTGCAGAAGATCTTTGGGAGCCTGGCATTTCTGCCGAGAGCTTTGAT	2876
Oy	329	-----	329
Db	2877	GGGAGCCAGCCTCCACACTGCCCCGCTCCAGCCAGAGACAGCTCCAAAGTGTAGACT	2936
Oy	329	-----	329
Db	2937	CTGAGAGAGATCAGACGTTACCTATACATCTCAGCATGGCCGGACAGCCCTGCCTGACCTC	2996
Oy	329	-----	329
Db	2997	AGCGTCTTCAGAACTGCAGATATTCGGGGAGACGAATTTCTGCACAAATGGCGCTACTCG	3056
Oy	329	-----	329
Db	3057	CTGACCTGCAGAGGGCTGGGCATCAGCTGGCTGGGGCTGCCTCAGCTAGAGGAAGTGGGC	3116
Oy	329	-----	329
Db	3117	AGTGGAGTGGCCCTCATCCACACAAATAAACCCACCTCTGCTTGTGTGCACACGGTGGCTGG	3176
Oy	329	-----	329
Db	3177	GACGAGCTTTTCGGAACCCGACCAAGACTGTCTCCACACTGCCAACCAGCGCAGAGAGAC	3236
Oy	329	-----	329
Db	3237	GAGTGTGGGCGAGGGCTTGGCTGCCACACAGCTGTGGCGCCGAGGGACACTGTGGGGT	3296
Oy	329	-----	329
Db	3297	CCAGGGCCACCCAGTGTGTCACTGCAGCCAGTCTCTTCGGGGCCAGAGATGCTGTGAG	3356
Oy	329	-----	329
Db	3357	GAATGCCGAGTACTGCAGAGGGCTTCCCAAGGAGTATGTAAATGCCAGAGCACTGTTTGGCG	3416
Oy	329	-----	329
Db	3417	TGCCACCTGAGTGTCAAGCCCAAGAAATGGCTAGTACCTGTTTTTGGACCCGAGGCTGAC	3476
Oy	329	-----	329
Db	3477	CAGTGTGTGGCCTGTGCCACATATAAAGACCTCCCTCTGCGTGGCCCGATGCCCAGC	3536
Oy	329	-----	329
Db	3537	GGTGTGAAACCTGACCTCTCTCATATGCCATCTGGAAAGTTTCCAGATAGAGAGGGCCCA	3596

329
3597 TGCCAGCCTTGGCCCATCACTGACCCACCTCTGTGTGACCTGATGACAAAGGCTGC 3656
329
3657 CCCGCCAGCAGAGAGCCAGCCCTCTGACGTCATGCTCTCTGGGTTGGCATTCCTG 3716
329
3717 CTGGTCTGGTCTTGGGGGTGGTCTTTGGATCTCATCAAGCGACGACAGAAATC 3776
329
3777 CGGAATACAGATCGGAGACTGCTCAGGAAAGAGCTGTGGAGCCGCTGACACT 3836
329
3837 AGCGAGCGATGCCCAACAGCGCGAGATCGGATCTTAAAGAGACGAGCTGAGGAG 3896
329
3897 GTGAAGTGTGGATGTGGGCTTTTGGCAGCTCAAGGCGATCTGGATCCCTGAT 3956
329
3957 GGGAGATGTGAATAATTCAGTGGCCATCAAGTGTGAGGGAAACACATCCCCAAA 4016
329
4017 GCCAACAAGAAATCTTAGACGAAGCATACGTGCTGCTGGTGGCTCCCATATGTC 4076
329
4077 TCCGCTCTGTGGCATCTGCTGACATCCAGGTGACGCTGTGACACAGCTTATGCC 4136
329
4137 TATGGCTCTTATGACCATGTCCGGGAAACCGCGAGCGCTGCCAGACCTG 4196
329
4197 CTGAATGTGTATGACGATTTGCCAAGGGATGAGTACTGAGAGATGTGGCTGCTA 4256
329
4257 CACAGGACTTGGCCGCTCGGAACGTGCTGTCAGAGATCCCAACATGTCAAAATTACA 4316
329
4317 GACTTCGGGCTGGCTGGCTGTGACATTTGACGAGACAGATACCATGAGATGGGGC 4376
329
4377 AAGTGGCCATCAAGTGGATGGCGCTGGAGTCCATTTCTCCGGCGGGTTCACCCACAG 4436
329
4437 AGTATGTGTGAGATGTATGTGTGACTGTGTGGAGCTGATGACTTTTGGGGCAACCT 4496
329
4497 TACGATGGATCCAGCCCGGAGATCCCTGACTGCTGAAAAAGGGGAGCGGCTGCC 4556
329
4557 CAGCCCCCATCTGCACCATTTGATGTCTACATGATGCTCAAAATGTGTGATGATTGAC 4616
329
4617 TCTGAATGTGGCCAAAGATTCCGGAGTGTGTCTGAATTTCCCGCATGGCCAGGAC 4676
329

4677 CCCAGCGCTTGTGTATCAAGATAGAGACTTGGGCCACGACCTTCCCTTGACAGC 4736
329
4737 ACCTTACCGCTCAGTCTGAGAGACGATGAGTGGGGACCTGTGTGATCTGAGAG 4796
330
4797 TATCTGATCCCGCAGAGGCTTCTTCTGTCACAGCCCTGCCCCGGCGGGGGCATG 4856
335
4857 GTCCACACAGGACCCGAGCTCATACAGAGTGGCGGTGGGAGCTGACACTGAGG 4916
355
4917 CTGAGCCCTGTGAAGAGAGGCCCCAGGCTCTCAGTCCAGCCCTCCGAAGGGGCTGGC 4976
375
4977 TCCGATGATTTGATGTGACCTGGAAATGGGGGACCAAGGGGCTGCAAGGCTGCC 5036
395
5037 ThrlsAspProSerProLeuGlnArgTyrSerGlnAspProThrValProLeuProSer 5096
415
5097 GAGACTATGCTACGTTGCCCCCTGACCTGACGACCCCGACGCTGATATGTAAACAG 5156
435
5157 ProAspValArgProGlnProProSerProArgGlnGlyProLeuProAlaAlaArgPro 5216
455
5217 AlaglyAlaThrLeuGlnAlaAlaLysThrLeuSerProGlyLysAsnGlyValLys 5276
475
5277 GCTGGTCCACTGTGAAGGGCCAAAGACTCTCCAGGAGAAATGGGCTGTCAA 5336
495
5337 AlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsnLeuTyr 5396
515
5397 TGGACCAAGACCCAGAGCGGGGGCTCCACCCAGCACTTCAAAAGGACACCTACG 5456
557
5457 GCAGAGAACCCAGAGTACCTGGGTCTGAGCTGCCA-----GTCGTAACCAAGAGCCA 5510
555
5511 SerProSerThrGlnProTyrGlnHisValAsnAlaIleGlnGlnAlaArgValLeu 5561
575
5562 AsnLeuSerArgAspThr 580
580
580259
4473 bp DNA linear PAT 31-AUG-2000
Sequence 1 from patent US 5968748.
AR080259
AR080259
AR080259.1 GI:10006994
Unknown.
Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 4473)
 AUTHORS Bennett, C. Frank, L. Lipton, A. and Witters, L. M.
 TITLE Antisense oligonucleotide modulation of human HER-2 expression
 JOURNAL Patent: US 5968748-A 1 19-OCT-1999;
 FEATURES Location/Qualifiers
 source 1..4473
 BASE COUNT 902 a 1383 c 1329 g 859 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,25e-80 Length: 4473
 Score: 2418.00 Matches: 523
 Percent Similarity: 41.06% Conserved: 5
 Best Local Similarity: 40.67% Mismatches: 22
 Query Match: 64.67% Indels: 736
 DB: Gaps: 4

US-09-821-883-2 (1-690) x AR080259 (1-4473)

QY 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
 Db 118 CGCCGCCCTCCAGCCGGGCTCCAGCCGACCATGGGGCCGGAGCCGACGTAGCACCATTG 177
 QY 18 PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
 Db 178 GAGCTGGGGGCGCTTGCGCGCTG-----GGGCTCTCTCGGCCCTCTTGCCCCC 228
 QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
 Db 229 GGAGCGCGCAGCACCACCAAGTGTGCACCGCGCAGACATGACCTCGCTCCCTGCCAT 288
 QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrgLysGluValValGlnGly 77
 Db 289 CCCGAGACCCACTGTGACATGCTCCGCGCACCTCTACCAAGGCTGGCCAGGTGTGCAGGGA 348
 QY 78 AsnLeuGluLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
 Db 349 AACCTGGAACTCACCCTACTGCGCCACCAATGCGAGCTCTCTCTCGAGATATCCAG 408
 QY 98 GluValGlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
 Db 409 GAGGTGCAGGGCTACGTCTCATCGCTCACAACCAAGTACAGCAGGTCCCTGACAGAG 468
 QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAsp 137
 Db 469 CTGGCGATTGTGCGAGGACCCAGCTCTTTGAGGACACTATGCCCCGCTGGCTAGAC 528
 QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
 Db 529 AATGGAGACCGGCTGACAAATACCAACCCCTGTACAGGGGCTCCCGCAGAGAGCCGTGG 588
 QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn 177
 Db 589 GAGCTGCACACTTCGAAGCTTCACAGATCTTGAAAGGAGGGGTCTTGATCCAGCCGAG 648
 QY 178 ProGlnLeuCysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln 197
 Db 649 CCCGACCTCTGCTACCGAGACACGATTTGTGAGAGACATCTTCCACAGAACCAACAG 708
 QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerPromet 217
 Db 709 CTGGGCTCTACACTGATATACACCAACCCCTCTCGGGCTGTGCACCCCTGTCTCGAAG 768
 QY 218 CysLysGlySerArgCysTrpLysLysSerGluAspCysGlnSerLeuThrArgThr 237
 Db 769 TGTAGAGGCTCCCGCTGCTGGGGAGAGAGTCTTGAGGATTTGACAGCCTGACGCGCAG 828
 QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu 257
 Db 829 GTCTGGGCGGGGTGCTGCTGCGCTGCAGAGGGCCACTGCCATGCTGCTGCGCCATTGAG 888
 QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277

Db 889 CAGTGTGCTGCCGGCTGCAGCGGGCCCAAGCACTGTGATGCTGCTGGCTCCACTTC 948
 QY 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThr 297
 Db 949 AACCAAGAGGGGATGTGAGGTGACATGCGCCAGCCCTGGTACCTACCAACACAGACAG 1008
 QY 298 PheGluSerMetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAla 317
 Db 1009 TTTGAGTCCATATCCCAATCCCAAGGCGGGTATACATTGCGGCCAGCTGTGACTGCC 1068
 QY 318 CysProTyrrAsnTyrrLeuSerThrAspValGlySer----- 329
 Db 1069 TGTCCCTACAACTACTTCTACGAGACGTGGATCTCTGACCCCTGTGCAC 1128
 QY 329 ----- 329
 Db 1129 AACCAAGAGTGACAGCAGAGATGGAACACAGCGGTGTGAGAACTGCAGAACGCCCTGT 1188
 QY 329 ----- 329
 Db 1189 GCCCGAGTGTGCTATGCTGTGGGCATGAGACACTTCCGAGAGGTGAGGCCACTTACAGT 1248
 QY 329 ----- 329
 Db 1249 GCCAATATCCAGAGATTGCTGCTGTGCAGAGAGATCTTTGGAGCCTGTGCTGCGC 1308
 QY 329 ----- 329
 Db 1309 GAGAGCTTTGATGGGAGCCACCCCTCCAACACTGCCGCCCTCCAGCAGCAGCTCCAA 1368
 QY 329 ----- 329
 Db 1369 GTGTTGAGACTCTGSAAGAGATCACAGGTTACTATACATCTCAGCATGCGCGACAGC 1428
 QY 329 ----- 329
 Db 1429 CTGCTGACCTCAGCGCTTTCCAGAACTGTCAAGTAATCCGGGAGCAATTTCCACAAAT 1488
 QY 329 ----- 329
 Db 1489 GCGGCTACTGCTGACCTCGAAGGGCTGGGCTGAGCTGCTGGGCTGCGCTCACTG 1548
 QY 329 ----- 329
 Db 1549 AAGGAACCTGGGAGTGGACTGGCCCTCATCCACATACACCCACTCTGCTTCTGTCAC 1608
 QY 329 ----- 329
 Db 1609 ACGGTGCCCTGGGACACGCTTTTGGAAACCGGACCAAGCTCTGTCCACACTGCCAAC 1668
 QY 329 ----- 329
 Db 1669 CGGCCAGAGACAGTGTGTGGGCGAGGGCTGGGCTGCCACACAGCTGTGCGCCGAGGG 1728
 QY 329 ----- 329
 Db 1729 CACTGCTGGGGTTCAGAGGCCCAACCAAGTGTTCAACTGACGACCAAGTTCTTGGGGCCAG 1788
 QY 329 ----- 329
 Db 1789 GAGTGGTGAGAGAAATGCCAGTACTGACAGGGGCTCCCGAGGAGATATGATGACAGAG 1848
 QY 329 ----- 329
 Db 1849 CACTTTTGGCGTGCACACCTAGTGTACAGCCCAAGAAATGGCTCACTGACTGTTTGA 1908
 QY 329 ----- 329
 Db 1909 CCGGAGGCTGACACAGTGTGTGCGCTGTGCCCACTATAGGACCTTCCTTGTGCGTGGCC 1968
 QY 329 ----- 329

Db 1969 CGCTGCCCAAGGCTGTGAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2028
QY 329 ----- 329
Db 2029 GAGGAGGGGGCATGCCAGCTTGCCCATCACTGCACCCACCTCTGTGGACCTGGAT 2088
QY 329 ----- 329
Db 2089 GACAAAGGCTGCCCCCGGAGAGAGAGCCAGCCCTTGCAGTCCATCATCTTGCGGTG 2148
QY 329 ----- 329
Db 2149 GTTGGCACTTCTGCTGCTGCTGCTTGGGGGTGGCTTTGGAGATCTTCATCAAGCAGG 2208
QY 329 ----- 329
Db 2209 CAGCAGAAAGATCGGGAAGTACAGATGCGAGACTGCTGAGGAAGAGAGCTGTGGAG 2268
QY 329 ----- 329
Db 2269 CCGCTGACACCTAGCGGAGCGATGCCCMACAGCGGAGATGCGGATCTGAAAGAGACG 2328
QY 329 ----- 329
Db 2329 GAGCTGAGGAAGGTGAAGGTGCTTGATCTGCGCTTTTGGCACAGTCTACAAAGGCATC 2388
QY 329 ----- 329
Db 2389 TGGATCCCTGATGGGAGAAATGTGAAATTCAGGTGCGCATCAAAAGTGTGAGGGAANA 2448
QY 329 ----- 329
Db 2449 ACAATCCCCCAAGCCAAAGAAATCTTAGACGAGACATACGTAGTGGTGTGGG 2508
QY 329 ----- 329
Db 2509 TCCCAATATGTCTCCGCTTCTGGGCACTGCTGACATCCAGCTGCAGCTGTGACA 2568
QY 329 ----- 329
Db 2569 CAGCTTATGCCCTATGGCTGCTCTTAGACCATGTCGGGAAAAACCCGAGCGCTGGG 2628
QY 329 ----- 329
Db 2629 TCCAGAGACTGCTGAAGCTGGTGTAGCAGATTGCCAAGGGAGTACCTACCTGGAGAT 2688
QY 329 ----- 329
Db 2689 GTGGGCTGTACACAGGAGCTTGGCCGCTCGGAACGTGCTGTCAAGAGTCCCAACCAT 2748
QY 329 ----- 329
Db 2749 GTCAAAATTACAGACTTCGGGCTGGCTGCGCTGTGACATTCAGACGAGACAGATACCAT 2808
QY 329 ----- 329
Db 2809 GCAGATGGGGGCAAGTGGCCCATCAAGTGAAGGCTGAGTCCATTCTCCGCGGG 2868
QY 329 ----- 329
Db 2869 TTCACCCACAGAGTATGTGTGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTT 2928
QY 329 ----- 329
Db 2929 GGGGCCAAACCTTACGATGGGATCCAGCCGGGAGATCCCTGACTGTGAAGAAAGGG 2988
QY 329 ----- 329
Db 2989 GAGCGGCTGCCAGCCGCCCATCTGCACCATTTGATGTCTACATGATGATGATCAATGT 3048
QY 329 ----- 329
Db 3049 TGGATATTGACTCTGAATGTGCGCCCAAGTTCGGGAGTTGGTGTCTGAATTCCTCCGC 3108

QY 329 ----- 329
Db 3109 ATGGCAGGGACCCCGAGCGTTTGTGTCATCCAGATGAGACTTGGGCCAGCCAGT 3168
QY 329 ----- 329
Db 3169 CCTTGGACAGACCTTCTACCGCTCAGCTGTGAGAGCATGACATGGGGGACCTGTG 3228
QY 330 ----- 330
Db 3229 GATGCTGAGAGTATCTGTAACCCAGAGGCTTCTTCTGTCCAGACCTGCCCGGCG 3288
QY 331 AAGllylmetValHisHisArgHisArgSerSerThrArgSerGlyGlyAsp 350
Db 3289 GCTGGGGGCAATGTCACACAGGACCGCAGCTCATCTACAGAGATGGGGGGGAGAC 3348
QY 351 LeuThrLeuGlyLeuGlyProSerGlyGlyAlaProArgSerProLeuAlaProSer 370
Db 3349 CTGACACTAGGGCTGGAGCCCTTGAAAGAGAGGCCCCAGGTTCTCCATGGCACCTCC 3408
QY 371 GluGlyAlaGlySerAspValPheAspGlyAspLeuGlyAlaLeuGlyLeu 390
Db 3409 GAAAGGCTGCTGCTCCGATGTTGATGTGATGCTGGGAATGGGGCCAGCCAGGCTG 3468
QY 391 GluSerLeuProThrHisAspProSerProLeuGlyAlaArgTyrSerGlyAspProThrVal 410
Db 3469 CAAAGCTCCCGACACATGACCCCGCCCTTACAGCGGTACAGTAGAGACCCACAGTA 3528
QY 411 ProLeuProSerGlyThrArgGlyTyrValAlaProLeuThrArgSerProGlyProLeu 430
Db 3529 CCGTCCCTCTGAGACTGATGCTACGTGCCCCCTGACCTGACCTGACCCCGAGCTGAA 3588
QY 431 TyrValAsnGlnProAspValArgProGlnProProSerProArgGlyGlyProLeuPro 450
Db 3589 TATGTGAACAGCAGATGTTCGGCCCGACCCCTTCCGCGAGAGGGCCCTGCTCCT 3648
QY 451 AlaAlaArgProAlaGlyAlaThrLeuGlyAlaAlaLysThrLeuSerProGlyLysAsn 470
Db 3649 GCTGCCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3708
QY 471 GlyValValLysAspValPheAlaPheGlyGlyValAlaGlyLysProGlyLeuThr 490
Db 3709 GGGCTGCTCAAGAGCTTTTTCCTTGGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3768
QY 491 ProGlnGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAsp 510
Db 3769 CCCAGGGAGAGAGCTGCCCTCAGCCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3828
QY 511 AsnLeuTyrTyrTrpAspGlnAspProProGlyAlaArgGlyAlaProProSerThrPheLys 530
Db 3829 AACCTTATTAATCTGGGACCGAGGCCACAGAGGGGGGCTCCACACGACCTTCAAA 3888
QY 531 GlyThrProThrAlaGlyAsnProGlyTyrLeuGlyLysAspValProAlaAlaAlaPro 550
Db 3889 GGCACACTTACGCGCAGAGAACCCAGAGTACCTGGGTCTGACGTGCCA-----GTGTGA 3942
QY 551 AlaArgSerProSerPro 556
Db 3943 ACCAGAAAGCCAAAGTCCG 3960
RESULT 13
ARI67390
LOCUS ARI67390 4473 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 26 from patent US 6287569.
ACCESSION ARI67390
VERSION ARI67390.1 GI:17903168
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4473)
AUTHORS Kipps,T.J. and Wu,Y.

TITLE	Vaccines with enhanced intracellular processing
JOURNAL	Patent: US 6287569-A 26 11-SEP-2001;
FEATURES	Location/Qualifiers
1	1173

BASE COUNT	902 a	1383 c	1329 g	859 t
ORIGIN				

Alignment Scores:

Pred. No.:	2,256-80	4472
Score:	2418.00	523
Percent Similarity:	41.06%	Conservative: 5
Best Local Similarity:	60.67%	Mismatches: 22
Query Match:	64.67%	Indels: 736
DB:	6	Gaps: 4

US-09-821-883-2 (1-690) x ARI67390 (1-4473)

OY		2	ArAlaIaPProLeuLeuLeuAlaAlaArgA-I-----AlaSerLeuLeuGly	17	
Db	118	CGRGCCCTCCACAGCGGGGTCCAGCCGGACCAATGGGCGGAGCCGACAGTGAAGCACCATG	118	CGCGCCCTCCACAGCGGGGTCCAGCCGGACCAATGGGCGGAGCCGACAGTGAAGCACCATG	177
OY		18	PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaIaArg	37	
Db	178	GAGTCGCAGCGGCTTGTCGCGGTGG- - - - -GGCTCTCCCGCGGCTTTGCCCCC	228	GGCTCTCCCGCGGCTTTGCCCCC	228
OY		38	GIYAAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer	57	
Db	229	GGAGCCCCGAGAACCCAAAGTGTGCACGGGACACACATGAAGCTGGCGCTCCGTCAGT	288	GGAGCCCCGAGAACCCAAAGTGTGCACGGGACACACATGAAGCTGGCGCTCCGTCAGT	288
OY		58	ProGluThrHisLeuAspMetLeuArgHisLeuLeuTrpGlnGlyCysGluValValGlnGly	77	
Db	289	CCCGAGACCCACCTGGAGCATCTCTGCCACACTCTACCAAGGGCTCCAGGTGGTCACAGGA	348	CCCGAGACCCACCTGGAGCATCTCTGCCACACTCTACCAAGGGCTCCAGGTGGTCACAGGA	348
OY		78	AsnLeuGluLeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97	
Db	349	AACCTGGAACCTACACTCTCTGCCACCAAATGCCAGCTGTCTCTCCGAGGATATCCAG	408	AACCTGGAACCTACACTCTCTGCCACCAAATGCCAGCTGTCTCTCCGAGGATATCCAG	408
OY		98	GluValGlnGlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117	
Db	409	GAGGTGACGGCTACGAGTCTATGCTGCACACAACCAATGAGGACGGTCCACCTGCACAGG	466	GAGGTGACGGCTACGAGTCTATGCTGCACACAACCAATGAGGACGGTCCACCTGCACAGG	466
OY		118	LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAsp	137	
Db	469	CTCGGGATTGTGGGAGGACACCCACTCTTTGAGSACAAATATGCCCTGGCGCTGTAC	528	CTCGGGATTGTGGGAGGACACCCACTCTTTGAGSACAAATATGCCCTGGCGCTGTAC	528
OY		138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157	
Db	529	AATGAGAACCCTCGAACAATATCACCCCTGTACAGGGGCTCCACAGAGGCTCGCG	588	AATGAGAACCCTCGAACAATATCACCCCTGTACAGGGGCTCCACAGAGGCTCGCG	588
OY		158	GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsn	177	
Db	589	GAGTGCAGCTCCAAACCTTCACAGAGATCTTGAAAGAGGGGCTTGATATCCACGGAAC	648	GAGTGCAGCTCCAAACCTTCACAGAGATCTTGAAAGAGGGGCTTGATATCCACGGAAC	648
OY		178	ProGlnLeuCysTrpGlnAspThrIleLeuThrLysAspIlePheHisLysAsnAsnGln	197	
Db	649	CCCAGCTGTGATCCAGACGACATTTTGTGGAAAGACATCTTCCACAAAGAACACAG	708	CCCAGCTGTGATCCAGACGACATTTTGTGGAAAGACATCTTCCACAAAGAACACAG	708
OY		198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet	217	
Db	709	CTGGCTCTCACACTGAATAGACACCAACCGCTCTGGGCTGGCCACCCTGTCTTCCGATG	768	CTGGCTCTCACACTGAATAGACACCAACCGCTCTGGGCTGGCCACCCTGTCTTCCGATG	768
OY		218	CysLysGlySerArgCysTrpGlyGlnSerSerGlnAspCysGlnSerLeuThrArgThr	237	
Db	769	TGTAAAGGCTCCGCTCTGGGAGGAAGATTCTGAGGATGTCTAGAACCTGTAGACCGACT	828	TGTAAAGGCTCCGCTCTGGGAGGAAGATTCTGAGGATGTCTAGAACCTGTAGACCGACT	828
OY		238	ValCysAlaGlyGlyCysAlaAlaArgCysLysGlyProLeuThrThrAspCysHisGlu	257	
Db	829	GTCGTGCGGATGTGCTGGCCGCTGCAAGGGGCACTGGCCAACTAGCTCTGCCATGAG	888	GTCGTGCGGATGTGCTGGCCGCTGCAAGGGGCACTGGCCAACTAGCTCTGCCATGAG	888
OY		258	GlnCysAlaAlaArgLysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277	
Db	889	CAGTGTCTCCGCTCTCACGGGCCCCAACACTATACCTGACCTGGCTGCTGCTCCACTTC	948	CAGTGTCTCCGCTCTCACGGGCCCCAACACTATACCTGACCTGGCTGCTGCTCCACTTC	948

QY	278	AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAsp	297
Db	949	AACCACAGTGGCATCTGTGTGAGTGCACCTGCCACGCCCTGTGATCCATACACACAGACACG	1008
QY	298	PheGluSerMetProAnProGluGluArgTyrThrPheGluAlaSerCysValThrAla	317
Db	1009	TTTAGTTCATGCCCATTCGCCATCCCGAGGGCCGGTATACATTCGGGGCCAGCGTGTACGTCC	1068
QY	318	CysProTyrAsnTyrLeuSerThrAspValGlySer-----	329
Db	1069	TGTCCCTACACACTACCTTTCTACGGAGCGTGGGTACTCTGCACCCTCGTGG;CCCTCGAC	1128
QY	329	-----	329
Db	1129	AACCAAGAGGTGACACGACGAGATGGAACACACAGCGGTGTGAGAAAGTGCACGACCCCTGT	1188
QY	329	-----	329
Db	1189	GCCCGAGTGTGATATGGTCTGTGGCATGAGACACTTGCAGAGAGTGGAGGAGATTACAGT	1248
QY	329	-----	329
Db	1249	GCCAAATATCCAGAGATTGTCTGCTGCACAGAAAGATCTTTGGAGCCTTGCAATTTCTGCCG	1308
QY	329	-----	329
Db	1309	GAGAGCTTTGATGGGGACCCAGCCTCCAAACACTGCCCCGCTCCAGCCAGAGCAGTCCAA	1368
QY	329	-----	329
Db	1369	GTGTTTGAAGCTGTGAAGAGATCACAGGTTACTATATCATCTCAGACTGCGCGACAGC	1428
QY	329	-----	329
Db	1429	CTGCCTGACCTCAGCGCTTCCAGAAACCTGCAAGTAACTCGGGGAGAGAAATTCGCACAT	1488
QY	329	-----	329
Db	1489	GGCGCTACTGTGACCTGCACCTGCAGAGGGCTGGGCATCAGCTGGGCTGGCTGCCCTACTG	1548
QY	329	-----	329
Db	1549	AGGGAACCTGGGCACTGAGTGGAGCTGCCCCCTCATCCACCAATACACCCACCTTGCTTGGTGAC	1608
QY	329	-----	329
Db	1609	ACGGTGCCCTGGGACACAGTCTTTCCGAACCCGACACCAAGCTCTGTGCTCAGACTGCCAAC	1668
QY	329	-----	329
Db	1669	CGGCCAGAGACGAGTGTGTGGGCGAGGGCTGGCCTGCACACAGCTGTGCCGCGAGAGG	1728
QY	329	-----	329
Db	1729	CACCTGCTGGGATCCAGGGCCACACCACTGTGTCACTGCAGCAGTTCCTTGGGGCCAG	1788
QY	329	-----	329
Db	1789	GAGTGGCTGGAGAAATGCCAGTACTGCAGAGGGCTCCCGAGGAGTATGTAAATGCCAGG	1848
QY	329	-----	329
Db	1849	CACCTGTTGCCGTGCCACCTGTAGTGTCAAGCCCAAGATGGCTCAGTACCTGTTTGGGA	1908
QY	329	-----	329
Db	1909	CCGAGAGCTGACACAGTGTGTGGCCTGTGCCACTATAAAGACCTGCCCTTGTGCTGGCC	1968
QY	329	-----	329
Db	1969	CGCTGCCCCACGGGTGAAACCTGCATCTCTCCATACATGCCCATCTGGAAGTTTCCAGAT	2028

OY	329	----	329	----	329
Db	2029	GAGGAGGGCGCATCCAGCCTTGCCCATCACTGCACCCACTCTCTGTGGACCTGGAT	2088	----	2088
OY	329	----	329	----	329
Db	2089	GACAGAAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTCTGACGTCATCATCTCTGGGGTG	2148	----	2148
OY	329	----	329	----	329
Db	2149	GTTGGCATTTCTGTGTCGTGCTTGGGGGGTGGTCTTTGGGATCCTCATCAAGGACGG	2208	----	2208
OY	329	----	329	----	329
Db	2209	CAGCAGAAGATCCGGAAATACAGATGCCGAGACTGCTGCAGGAACGGAGCTGGTGAG	2268	----	2268
OY	329	----	329	----	329
Db	2269	CCGCTGACACCTAGCGAGGAGATGCCCAACGAGCGCAGATGGCGATCTGAAAGAGAG	2328	----	2328
OY	329	----	329	----	329
Db	2329	GAGCTGAGGAAGTGGAAGTGCTTGGATCTGCGGCTTTTGGCACAGCTTACAAGGACATC	2388	----	2388
OY	329	----	329	----	329
Db	2389	TGGATCCCTGATGGGGGAATGTGMAAATTCCAGTGGCCATCAAACTGTTGAGGGAAAC	2448	----	2448
OY	329	----	329	----	329
Db	2449	ACATCCCCCAAGCCACAAAGAAATCTTAGCAGGAAGCATAGCTGATGGCTGATGGGC	2508	----	2508
OY	329	----	329	----	329
Db	2509	TCCCATATGTCTCCGCGCTTCTGGGCATCTGCCTGACATCCACGGTGAGCTGTGACA	2568	----	2568
OY	329	----	329	----	329
Db	2569	CAGCTATGCCCTATGGCTGCTCTTTAGACCAATGTCCGGAAAAACCGGAGCCTGGGC	2628	----	2628
OY	329	----	329	----	329
Db	2629	TCCCAGACCTGCTGAACCTGATGATGCAGATTGCCAAGGGATGAGCACTACCTGGAGAT	2688	----	2688
OY	329	----	329	----	329
Db	2689	GTCGGCTCTGACACAGGAGACTTGGCGCTCGGAACGTGCTGATCAAGATCCCAACAT	2748	----	2748
OY	329	----	329	----	329
Db	2749	GTCMAAATTACAGACTTCGGGCTGGCTCGGCTGCTGACATTGACGAGACAGATACAT	2808	----	2808
OY	329	----	329	----	329
Db	2809	GCAGATGGGGSCAAGGTGCCCATCAAGTGAATGGCGCTGAGTCCATTCTCCGCGGGCG	2868	----	2868
OY	329	----	329	----	329
Db	2869	TTCAACCAACAGACTGATGTGGAGTTATGTGTGACTGTGTGGAGCTGATGACTTTT	2928	----	2928
OY	329	----	329	----	329
Db	2929	GGGGCCAAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTGCTGGAAAAAGGG	2988	----	2988
OY	329	----	329	----	329
Db	2989	GAGCGGCTGCCCCAGCCCCCATCTGCACCATTGATGTCTACATGATCATGTGCTCAAATGT	3048	----	3048
OY	329	----	329	----	329
Db	3049	TGGATGATGACTCTGAATGTGGGCCAAAGATTCCGGGAGTTGGTGTCTGAATTTCCCGC	3108	----	3108
OY	329	----	329	----	329

Db	Accession	Source	LOCUS	DEFINITION	VERSION	KEYWORDS	ORGANISM	REFERENCE
Db	3109	ATGGCCAGGAGACCCCCAGGGCTTTGTGGTCATCCAGAAATGAGAGCTTGGGCCACGACAGT	3168					
QY	329	-----	329					
Db	3169	CCCTTGACACGACCTTACCCTACCTGCTGAGAGACGATGACATGGGGACCTGGTG	3228					
QY	330	-----	330					
Db	3229	CATGCTGAGAGATATCTGGTACCCGACGAGGGCTTCTTCTGTCGACAGCCCTGCCCGGGC	3288					
QY	331	AlAGlYglYmeVAlhHsHsArghHsArghSerSerSerThrArgSerGlyGlyGlyAsp	350					
Db	3289	GCTGGGGCATGTCGACACACGACGACGACGACCTCATCTACAGAGATGGCGGTGGGAC	3348					
QY	351	LeuThrIleuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSer	370					
Db	3349	CTGACACTAGAGGCTGGAGCCCTGTAAAGAGGGGCCCCAGGCTCTTCACTGGCACCCCTCC	3408					
QY	371	GlUGlYAlAGlYSerAspValPheAspGlyAspLeuGlyMeGlyAlaAlaGlyGlyLeu	390					
Db	3409	GAAAGGGCTGGCTCCGATGATATTATGTGATGACCTGGGAATGGGGGACCCAGAGGGCTG	3468					
QY	391	GlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrVal	410					
Db	3469	CAAAAGCTCCCAACATGACACCCAGCCCTTACAGCGGTACAGTACAGAGACCCCACTA	3528					
QY	411	ProLeuProSerGluThAspGlyTyrValAlaProLeuThrCysSerProGlnProGlu	430					
Db	3529	CCCTGCGCTCTGAGACTGATGGCTACGTTGCCCTGACCTGACGCCCCCAAGCTTGA	3588					
QY	431	TyrValaIngInProAspValArgProGlnProProSerProArgGluGlyProLeuPro	450					
Db	3589	TATGTGAACACACAGATGTTGGGCCCAAGCCCTTGCCTGCGCCGAGAGGGCCCTGTGCT	3648					
QY	451	AlaAlaArgProAlaGlyAlaThrIleuGluArgAlaLysThrLeuSerProGlyLysAsn	470					
Db	3649	GCTGCCACACCTCGCTGGTGCCTGTGAAAGGCCCAAGACTCTCCCAAGGGAAGAT	3708					
QY	471	GlyValValLysAspValPheAlaPheGlyGlyAlaValaGluAsnProGluTyrLeuThr	490					
Db	3709	GGGGTCGTCAAGACAGCTTTTGTGCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACA	3768					
QY	491	ProGlnGlyGlyAlaAlaProGlnProThrHisProProProAlaPheSerProAlaPheAsp	510					
Db	3769	CCCCAGGGAGGAGTCCCTCAGCCCAACCTCTCTGCTTCAAGCCAGCCCTTGCAC	3828					
QY	511	AsnLeuTyrTyrTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLys	530					
Db	3829	AACCTCTATTACGTGGACAGACACACACAGCGGGGGCTCCACCCAGACACCTTCAAA	3888					
QY	531	GlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaPro	550					
Db	3889	GGGACACCTACGCGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCA-----GTGTGA	3942					
QY	551	AlaArgSerProSerPro	556					
Db	3943	ACCAGAGGCCCAAGTCCG	3960					
RESULT 14								
HSERB2R								
LOCUS	DEFINITION	4473 bp	Linear	PRI 30-MAR-1995				
ACCESSION	X03363							
VERSION	X03363.1	GI:31197						
KEYWORDS	cell surface glycoprotein; cellular oncogene; erb-2 cellular; glycoprotein; growth factor receptor; kinase; neu cellular oncogene; transmembrane protein; tyrosine kinase.							
SOURCE	Homo sapiens.							
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.							
REFERENCE	1 (bases 1 to 4473)							

FEATURES	source
AUTHORS	Yamamoto, T., Ikawa, S., Akiyama, T., Sema, K., Nomura, N., Miyajima, N., Saito, F. and Toyoshima, K.
TITLE	Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor
JOURNAL	Nature 319 (6050), 230-234 (1986)
MEDLINE	86118663
PUBMED	3003577
REFERENCE	2 (bases 1 to 4473)
AUTHORS	Papewalis, J., Nikitin, A.Yu. and Rajewsky, M.F.
TITLE	G to A polymorphism at amino acid codon 653 of the human erbB-2/Her2 gene
JOURNAL	Nucleic Acids Res. 19 (19), 5452 (1991)
MEDLINE	92020265
PUBMED	1681519
COMMENT	The c-erb-B-2 protein shows similarity to the epidermal growth factor receptor.
FEATURES	Location/Qualifiers
source	1..4473
	/organism="Homo sapiens"
	/strain="cell line MKN-7"
	/db_xref="taxon:9606"
	175..3942
CDS	/note="c-erb-B-2 precursor"
	/codon_start=1
	/protein_id="CAA27060.1"
	/db_xref="GI:31198"
	/db_xref="SWISS-PROT:P04626"
	/translation="MELALRCRMGLLALLALPRGASNVQVCGTDMRLRASPETHDMLRHLYOGCCVQVGNLELYLPTNASSLSFDIDIOVGYLIANNQRYPLRLRYRGTDLELMDALVLDGDDPLNNTPTPTGASPGGLRELDLSLEILKGLVLRNRPQLCYDITLMDIDFHKNNOLALITIDTNRACRHCSPKSGSRGWSSEDCOSTLRPVCAAGCARCKGRLPTDCCHEGCAAGCTCPKHSIDLACLFHNSGICELHOPALVTNTDTFSMPNPEGRYTFGASCVACPYANLSTDSVGSCTLPCHNQEVTAEDGTORCEKSKPCARVQYEGMEHLREVRVTSANTIOEFAGCKIKGSLAFLEPSEGDPASTADLPQEOLOYETLEETLYGLYISAMPDLSLDLSVFNQLVIGRLILHNGAYSLTLOGGILSMGLRSLRELSGIALIHNTNHCAPYHVPMDLFENPHALLHTNRPDEVGSLACIQLCARHCGRGPTQCYNCSGFTQGCVECRVLGPREIYNARHCLPHEFCOPNGSVTCGPEPADOCVAANH KDDPPVCARPSVKKDLSLTPMKRPDEBGACQPCPLNCHSCVDLDKCGPAEQASPLTSLISAVGILVVLGVGILIKRQOKT RKYMRRLLOETLELYLPTESGAMPNOAMRLKETELRKVLVLSGAFGVYKGIIPDGNNVLPVAILKYLRENTSPKANKLELIDEXYVAGVSPVSRSLGICLSTVOLVTLQMLVEGCLLDHVRNRRGRISODLLNMCOMAKMSYLEDVRLVHRLAARNVYKSPNHYKTPDGLARLIDIDETEVHADGQVPIKMALESILRRPFRHSDNYSCTYVLEMTGARAPYDGIPTRELPLEKGERLPQPICTITDYMIMVKMIDSECRFRRELVTSESRNARDQRFVVIQNEDEDGASPLDSTFYRLLEDDMDGLVDAEELVQDGEFCPPRPAAGAGGVHNRHRSSTRSGGDLITGLPSESEARSPSLASSEGSGVDFGFCIDGSAKGLSLPTHPSPLOKQSEDPVLPVLEETDGYVALPLTCSQPEVYNQDVPAPQSPRSGPLPARPAGATLERKTLSPKNGVGVKVFAPGAVENPEYTLTPGGAAAPQHPAPSPAPFNLYYWMDDPRECAPSTFNGCTPLAENPELGLDIVY"
misc_peptide	175..237
	/note="putative"
	238..3939
misc_feature	376..384
	/product="put. c-erb-B-2 protein (aa 1-1234)"
misc_feature	544..558
	/note="pot. glycosylation site"
misc_feature	733..741
	/note="pot. glycosylation site"
misc_feature	949..957
	/note="pot. glycosylation site"
misc_feature	1762..1770
	/note="pot. glycosylation site"
misc_feature	1885..1893
	/note="pot. glycosylation site"
misc_feature	2059..2067
	/note="pot. glycosylation site"
misc_feature	2353..3132
	/note="aa 727-986, seq. homologous to EGF receptor kinase domain"
misc_feature	2440..2454
	/note="pot. glycosylation site"
misc_feature	4455..4460

polyA_site		/note="put. polyA signal"	
4473		4473	
BASE COUNT		902 a 1383 c 1329 g 859 t	
ORIGIN			
Alignment Scores:			
-Pred. No.:		2.25e-80	
Score:		2418.00	
Percent Similarity:		41.06%	
Best Local Similarity:		40.67%	
Query Match:		64.67%	
DB:		9	
US-09-821-883-2 (1-690) x HSERR2R (1-4473)			
QY	2	ArglaAlaIaPLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly	17
DB	118	CGCGCCCTCCACAGCGCGGTCCAGCGCGGCATGGGGCCGAGCGGAGTGCACCATG	177
QY	18	PhelLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg	37
DB	178	GAGCTGGGGGCTTGTGGCGCTGG-----GGCTCTCTCTCGCCCTCTTCCGCC	228
QY	38	GlyAlaAlaSerThrGlnValAlcysThrGlyThrAspMetLysLeuArgLeuProAlaSer	57
DB	229	GGAGCGCGAGACCCAGCATGTGCAACGGCAGACACATGAAGCTGCGGCTCCCTCCAGT	288
QY	58	ProGluThrHisLeuAspMetLeuArgHisLeuArgGlnGlyCysGlnValAlcGlnGly	77
DB	289	CCCGAGACCCACCTGGACATCTCTGGCCACCTTACCAAGGCTCCAGTGTGTCCAGGA	348
QY	78	AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
DB	349	AACCTGAACCTACCTACTCTGCGCCACCAATCCAGACCTGTCTCTTCAGATATCCAG	408
QY	98	GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
DB	409	GAGGTGACGGGTGTACTCTCATGCTTCACAAACCAAGAGGAGGCTCCACTGCAGAGG	466
QY	118	LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp	137
DB	469	CTGGCGGATTTGTGCGAGGACCCAGCATCTTTAGAGCAACATATGCGCTGCGCTGAC	528
QY	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
DB	529	AATGAGACACCGCTGACAAATACCAACCCCTCTCAAGGGCTCCCGAGGAGGCTCCGG	586
QY	158	GluLeuGlnLeuArgSerLeuThrGlnIleLeuLysGlyGlyValLeuIleGlnArgAsn	177
DB	589	GAGCTGCAGCTTCGAAAGCTCACAGATCTTTGAAGAAGAGGGCTTGTATCCAGCGAAC	648
QY	178	ProGlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGln	197
DB	649	CCCCAGCTCTGCTACAGAGACACGATTTTGTGGAAGCAATCTTCCACAGAACACACAG	708
QY	198	LeuAlaLeuThrLeuIleAspThrAspArgSerArgAlaCysHisProCysSerProMet	211
DB	709	CTGGCTCTACACTATAGACACCAACGCTCTGGGGCTGTCCACCCCTGTCTCCAGATG	766
QY	218	CysLysGlySerArgCysTyrGlyLcyluSerSerGluAspCysGlnSerLeuThrArgThr	237
DB	769	TGTAAAGGCTCCCGCTGCTGTGGGAGAGAGTTCTGAGATTTCTCAGAGCTGACCGCACT	828
QY	238	ValGlyAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrArgCysCysHisGlu	257
DB	829	GCTCTGTGCGGGGTGTGGCCCGGTGCAAGGGGCCACATGACTGTGTGCATGAG	888
QY	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277
DB	889	CAGTGTCTGCTCGGCTGTGCACGGGCGCCAAAGCACTGTAGTCTCTGCGCTGCTCACTTC	948
QY	278	AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297

Dd	949	AACCAAGTGGCATCTGTGAGACTGCACCTGCCCGCCCTGGTGACATAACAACAGACAGC	1008
Oy	298	PheGIUserMetProAsnProGluIleArgTyrrThPheGlyAlaSerCysValIhrIala	317
Dd	1009	TTTGAGTTCATGCCCCAATCCCAGAGGGCGGTAAATATTGGCGCCAGCTGTGTGACTGCC	1066
Oy	318	CysProTyraAsnTyrlLeuSerThrAspValGlySer	329
Dd	1069	TGTCCCTAACACTACCTCTTCTACGGAGCTGGGATCCTGGACCCTGCTCTGCCCCCTGCAC	1122
Oy	329	-----	329
Dd	1129	AACCAAGAGTGAACAGACAGAGATGAAACACAGCGGTGTGAGAAGTGACAAAGCCCTGT	1188
Oy	329	-----	329
Dd	1189	GCCCGAGTGTCTATGTGCTGTGGCANTGAGCACCTTGCAGAGGTAGGGCAGTTACCACT	1248
Oy	329	-----	329
Dd	1249	GCCAATATCAGAGAGTTTGCTGGCTGCANAGAAATCTTTGGAGACCTTGGCATTTCTGCCG	1308
Oy	329	-----	329
Dd	1309	GAGACCTTTGATGGGGAGCCAGCCTCCAMACATGCCCGCTCCAGCCAGAGCAGCTCCAA	1368
Oy	329	-----	329
Dd	1369	GTTGTTGAGACTCTTGAAAGAGATCACAGGTTTACCTTATACATCTACAGATGSCCGAGACGC	1428
Oy	329	-----	329
Dd	1429	CTGCCTGACCTCAGACGTTCTTCCAGAAACCTTCAGATTAATCGGGGAGAGATTCTGCACAAT	1488
Oy	329	-----	329
Dd	1489	GGCGCTACTGCTGACACCTTGCAAAGGCTGGCATCAGCTGGCTGGGGCTGCGCTACAG	1548
Oy	329	-----	329
Dd	1549	AGGGAAGTGGGCACTGGACATGCGCCCTCATCCACCATTACACCCACCTCTGCTTGGTGAC	1608
Oy	329	-----	329
Dd	1609	ACGGTCCCTGGGACACAGCTTTTCGAAACCCGACCAAGCTGTGCTCCACACTGCCAAC	1668
Oy	329	-----	329
Dd	1669	CGGCCAGAGACGAGTGTGTGGGAGAGGCTTGCTGCCACCAGCTGTGCCCGCAGAGG	1728
Oy	329	-----	329
Dd	1729	CACGTGCTGGGATCCAGGGCCOACOCAGTGTCACTGACGACCAAGTTCTTGGGGCCAG	1788
Oy	329	-----	329
Dd	1789	GAGTGCCTGAGAGAAATCCGAGTACTGACAGGGGCTCCGAGGAGTATGTGAATGCCAGG	1848
Oy	329	-----	329
Dd	1849	CACGTGTTGGCCGTCCACCTCGAGTGCAGCCCCAGAAATGGCTCAGTAGCACTGTTTTGGA	1908
Oy	329	-----	329
Dd	1909	CCGGAGGCTGACAGTGTGTGGCTGTGCCCACTATAAGGACCCTCCCTTCTGCGTGGCC	1968
Oy	329	-----	329
Dd	1969	CGCTGCCCAAGCGGTGTGAAACCTGACCTCTACATGCCATCTGAAAGTTTCCAGAT	2028
Oy	329	-----	329

Db	2029	GAGGAGGGCGCATCGACGCTTCCCATTCACACTGCACCCACTCTCTGTGTAGCACTGGAT	2088
QY	329	-----	329
Db	2089	GACAGAGGCTGCCCGCCGAGCAGAGAGCCAGCCCTTGACGTCCATCATCTTCTCGGGTG	2148
QY	329	-----	329
Db	2149	GTTGGCATTTGCTGTCGTCGTCGTCCTTGGGGGTCGTCTTTGGGATCCTCATCAGCAGCG	2208
QY	329	-----	329
Db	2209	CAGCAGAGATCCGGAAGTACAGATGCGAGACTGTCAGAGAAAGAGCTGTGAG	2268
QY	329	-----	329
Db	2269	CCGCTGACACCTACGGAGGATGCCCAACAGGGCGCAATGCGGATCTCTAAGAGACG	2328
QY	329	-----	329
Db	2329	GAGGTAGAGAGTGAAGGTCTTGGATCTGCGCTTTTGGCACAGTCTACAGAGGCAATC	2388
QY	329	-----	329
Db	2389	TGATCCCTATGGGGAGAAATGTAAAAATTCCAGTGGCATCAAAAGTTGAGGSAAAC	2448
QY	329	-----	329
Db	2449	AACATCCCAAGCCCAACAAAGAAATCTTAGACGAAGATACGTAGTGAGCTGTGTGGC	2508
QY	329	-----	329
Db	2509	TCCCAATATGTCCTCCGCTTCTGGGCATCTGCTGACATCCAGGTCGACGTGTGACA	2568
QY	329	-----	329
Db	2569	CAGCTATGCCCTATGGCTGCTCTTAAACATGTCCGGGAAACCGGAGCCTGGGC	2628
QY	329	-----	329
Db	2629	TCCAGAGACTGCTGAACGTGTATGCAGATTGCCAAGGGATGAGTACCTGAGAGAT	2688
QY	329	-----	329
Db	2689	GTCGGGCTCGTACAGAGGACTTGGCCGCTCGAAGCTGCTGTAAGATGCCAACAT	2748
QY	329	-----	329
Db	2749	GTCAAAATTACAGACTTCGGGCTGCGCTCGCTGTGACATTGACAGACAGAGTACCAT	2808
QY	329	-----	329
Db	2809	GCAATGGGGCAAGTGCCCATCAAGTGAATGGGCTGAGTCCATTCTCGCCGGCGG	2868
QY	329	-----	329
Db	2869	TTCAACCACAGACTGATGTGAGATTATGTCGACTGTGTGGAGCTGATGACTTTT	2928
QY	329	-----	329
Db	2929	GGGGCCAAACCTTACGATGGGATCCAGCCCGGAGATCCCTGACCTCTGTAAAGGGG	2988
QY	329	-----	329
Db	2989	GAGGGGCTGCCCCAGCCCCCATCTGCACATTGATGTCTACATGATCATGTGTAATGT	3048
QY	329	-----	329
Db	3049	TGATGATGACTCTGAATGTGCGCCAAATTCGGGAGATTGGTCTGAATTCTCCGC	3108
QY	329	-----	329
Db	3109	ATGGCCAGGACCCCCAGCGCTTTGTGTGTCATCCAGATGAGAGCTTGGGCTCCAGCCAGT	3168

[illegible]

FEATURES		Genentech, Inc. (US)		Location/Qualifiers	
SOURCE	1-3768	/organism="Homo sapiens"			
BASE COUNT	758 a 1170 c 1121 g 719 t	/db_xref="taxon:9606"			
ORIGIN					
Alignment Scores:					
Pred. No.:	3 59e-80	Length:	3768		
Score:	2410.50	Matches:	510		
Percent Similarity:	41.16%	Conservative:	0		
Best Local Similarity:	41.16%	Mismatches:	2		
Query Match:	64.47%	Indels:	727		
DB:	6	Gaps:	1		
US-09-821-883-2 (1-650) x AX060704 (1-3768)					
QY	35	LeuAlaArgIyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeu	54		
DB	46	TTGGCCCCCGAGCGCGGAGACCCCAAGTGTGCACCGGACAGACATGAAGCTCGGGCTC	1059		
QY	55	ProAlaSerProGluThrHisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnVal	74		
DB	106	CTGCGCACTCCCGAAGCCACCTGGACATGCTCCGCCACCTTACCAAGGCTGCCAGGTG	1655		
QY	75	ValGlnGlyAsnLeuGlnLeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGln	94		
DB	166	GTGCAGGGAACCTGGACATCCACTACCTGGCCACCAATGCCACGCTGCTCTGTCAG	2259		
QY	95	AspIleGlnGluValGlnGlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValPro	1144		
DB	226	GATATCCAGAGAGCTCCAGGCTAGCTGTCTCATGCTCTCAACCAAGTGAAGGACAGTCCA	2895		
QY	115	LeuGlnATrGLeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAla	1344		
DB	286	CTGCAGAGGCTGGGGATTGTGTGCAGAGCCACGCTCTTTGAGCAACATAAGCCCTGGCC	3459		
QY	135	ValLeuAspAsnGlyAspProLeuAsnAsnThrThrProValIlnrGlyAlaSerProGly	1549		
DB	346	GTGTCTAGACAAATGGAGAGCCCTGTAAACAAATGCACCCCTGTCAACAGGGGCTCCCAAGA	4059		
QY	155	GlyLeuATrGgluLeuGlnLeuArgSerLeuThrGlnIleLeuLysGlyValIleIle	1749		
DB	406	GGCTCTCGGGAGACTCCACTTCGAAAGCTTCACAGATCTTGAAGAGAGGGGCTTTATATC	4659		
QY	175	GlnATrAspAsnProGlnLeuCysTyrrGlnAspThrIleLeuTyrrLysAspIlePheHisLys	1949		
DB	466	CAGCGGAACCCCCACTCTGTCTACAGAGACAGATTTGTGTGAAGAGCATCTTCCACAG	5259		
QY	195	AsnAsnGlnLeuAlaLeuThrLeuIleAspThrAsnArgSerATrGlyAlaCysHisProCys	2149		
DB	526	AACAACACAGCTGGCTCTACACTGTATACACCAACCGCTCTCGGGCTGTGCACCTGT	5899		
QY	215	SerProMetCysLysGlySerArgCysTyrrPylGluSerSerGlnAspCysGlnSerLeu	2349		
DB	586	TCTCCGATGTGTGAAGGCTCCCGCTGCTGGGGAGAGATTTGTAGAGATGTGCACAGCTG	6459		
QY	235	ThrATrThrValCysAlaGlyLysAlaArgCysLysGlyProLeuProThrAspCys	2549		
DB	646	ACGGCACTGTCTGTGGCGGGTGGCTGTCCCGGTGCAAGGGGGCCACTGCGCCACTGACTGC	7059		
QY	255	CysHisGlnGlnCysAlaAlaGlyCysThrGlyProLysHisSerATrAspCysLeuAlaCys	2749		
DB	706	TGCCATGTAGCAAGTGTGTGGCGGGTGGACAGGGGCCCAAGCACTGTGACTGTGTGGCTGTG	7659		
QY	275	LeuHisPheAsnHisSerGlyIleCysGlnLeuHisCysProAlaLeuValThrTyrrAsn	2949		
DB	766	CTCCCACTTCAACCAAGTGGCATGTGTAGCTGTGACAGTCCAGGCTGTGTCACTAC	8259		
QY	295	ThrAspThrPheGluSerMetProAsnProGluLysATrGlyTyrrPheGlyAlaSerCys	3149		
DB	826	ACAGACAGCTTTGATTCATCCCAATCCCAAGGGCGGGTATATTCGGGGCAGAGCTGT	8859		

Oy	315	ValThralAcysProTyraAsnTyrlLeuSerThraValGlySer-----	329
Db	886	GTGACTGCTGTCTCCCTACAACTACCTTCTTACGAGAGTGGATCTGCACCCCTCGTTCGC	945
Oy	329	-----	329
Db	946	CCCCGCACAAACAAAGAGGTGACAGCAGAGATGGAACAACAGCGGTGTGAGAAATGCACG	1005
Oy	329	-----	329
Db	1006	AAGCCCTGTGCCGAGGTGTGCTATGTGTGTGGCATGAGACACTTGGCAGAGGTGAGGGCA	1065
Oy	329	-----	329
Db	1086	GTTACAGTGGCCAAATATCCAGAGATTTCGTGGCTGCAGAAAGATCTTTGGAGCCTGGCA	1122
Oy	329	-----	329
Db	1126	TTTGTCCGGAGAGCTTTGTATGGGGACCCAGCCTCCAAACTGCCCGCTCCAGCCAGAG	1185
Oy	329	-----	329
Db	1186	CAGCTCCAAGTGTGTGAGACTCTGGAAGAGATCAGAGTTACCTATACCTTCACAGATG	1245
Oy	329	-----	329
Db	1246	CCGGACAGCCTGCTGACCTCAGCGTCTTCACAAACCTCAAGTAATCCGGGGAGCAATT	1305
Oy	329	-----	329
Db	1306	CTGCACAATGGCCCTACTCGCTGACCTTGCAAGGCGTGGGCAATCAGCTGGCTGGGCTG	1365
Oy	329	-----	329
Db	1366	CGCTACTGAGAGGAATCTGGGCGAGTGGAGCTGGCCCTCATCACCAATPAACACCACTCTGC	1425
Oy	329	-----	329
Db	1426	TTCTGTACACGAGTGCCCTGGGACCAAGCTTTTCGGAACCCGACCAAGCTCTGCTGCAC	1485
Oy	329	-----	329
Db	1486	ACTGCCAAACGGGCCAAGAGAGCAAGATGTGTGGGGAGGGCCTGGCTGCCACAGCCTGTC	1545
Oy	329	-----	329
Db	1546	GCCGAGGGGACTGTGGGGTCCAGGGCCACCACCACTGTGTCAACTGCACGACATTCCTT	1605
Oy	329	-----	329
Db	1606	CGGGGCCAAGAGTGTGTGGAGGAATGCCAGTCTGTCAAGGGGCTCCCCAGGAGTATGTG	1665
Oy	329	-----	329
Db	1666	AATGCCAAGCACTGTTCGCCGTGCACACCTTGAGTGTACGCCCAAGAAATGTCAGTGAAC	1725
Oy	329	-----	329
Db	1726	TGTTTGGACCGAGGCTGCACAGTGTGTGGCTGTGCCACTATTAAGACCCCTCCCTTC	1785
Oy	329	-----	329
Db	1786	TGCGTGGCCCGCTGCCACCGGTGTGAAACCTGCACCTCTCTACATGCCCATCTTGGAA	1845
Oy	329	-----	329
Db	1846	TTTCCAGATGAGAGGGCGCATGCCAAGCCTTGCCCCATCAACTGCACCCACTCTGTGTG	1905
Oy	329	-----	329
Db	1906	GACCTGATGACAAGGCTGCCCGCCCGAGACAGAGACCAAGCCGCTCTGACGTCCATCGTC	1965

QY	329	-----	329
Db	1966	TCTGCGGTGGATTGCAATTCGTGCTGCTGCTTGCGGGGTGCTTTGGATCCTCATC	2025
QY	329	-----	329
Db	2026	AAGGACGGCAGCAGAAAGATCCGGAAGTACAGATGCCGAGACTGCTGCAGAAACGGAG	2085
QY	329	-----	329
Db	2086	CTGGTGAGCCGTGACACCTAACGGAGGAGTCCCAACAGGCGCAGATGGGATCCTG	2145
QY	329	-----	329
Db	2146	AAAGAGACGGAGCTGAGAGAGTGAAAGTGCTTGATCTGGCGCTTTTGGACAGTCTAC	2205
QY	329	-----	329
Db	2206	AAGGCACTTGATCCCTGATGGGAGAAATGTGAAAATTCAGTGGCATCAAGTGTG	2265
QY	329	-----	329
Db	2266	AGGGAACACATCCCAACAGCCAAAGAAATCTTAGACGAAGCATACGTATGGCT	2325
QY	329	-----	329
Db	2326	GATGTGGCTCCCATATGTCTCCGCTTCTGGGCATTCGCTGACATCCACGCTGACG	2385
QY	329	-----	329
Db	2386	CTGGTGACACAGCTTATGCCCTATGGCTGCTCTTAGACCATGTCCGGAAAACTCGCGA	2445
QY	329	-----	329
Db	2446	CGCCTGGCTCCAGACACTGTGAACGTGTGATGACAGATTGCCAAGGAGTAGCTAC	2505
QY	329	-----	329
Db	2506	CTGGAGATGTGGCGCTCGTACACAGGACCTTGGCCGCTCGGAACGTCTGTCAAGACT	2565
QY	329	-----	329
Db	2566	CCCAACCATGTCAAAATTTACAGACTTGGGGTGGGCTCGGCTGCTGGACATTGACAGACA	2625
QY	329	-----	329
Db	2626	GAGTACATGACAGATGGGGGCAAGTGCCCATCAAGTGAATGGCGCTGAGTCCATTCTC	2685
QY	329	-----	329
Db	2686	CGCCGGCGGTTCAACCCACAGAGTGAATGTGTGAGTTATGGTGTACTGTGTGGAGCTG	2745
QY	329	-----	329
Db	2746	ATGACTTTTGGGCGCAAACTTACAGATGGGATCCAGCCCGGAGATCCCTGACCTGCTG	2805
QY	329	-----	329
Db	2806	GAAGAGGGGAGCGGCTGCCCAAGCCCCCATCTGCACCATTTGATGTCTACATGATCATG	2865
QY	329	-----	329
Db	2866	GTCAAAATGTTGATGATTGACTGTGAATGTGCGCCAAAGTTCCGGAGATTGTTGTTGAA	2925
QY	329	-----	329
Db	2926	TTCTCCGCATGGCCAGGACCCCAAGCGCTTGTGTGTCATCAGAAATGAGGACTTGGGC	2985
QY	329	-----	329
Db	2986	CCAGCCAGTCCCTTGGACAGACACTTCTACCGCTCACTGCTGGAGAGAGATGACATGGGG	3045
QY	329	-----	329

Db 3046 GACCTGTGATGCTGAGAGACTATCTGTACCCAGAGGGCTTCTTCTGTCCAGACCT 3105
QY 330 -----GlyAlaGlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGly 347
Db 3106 GCGCGGGGCTGGGTGGATGTCACACAGGACCGCAGCTCATCTACAGGAGTGCC 3165
QY 348 GlyAlaSerLeuThrLeuGlyLeuGlyLeuProSerGluGluAlaProArgSerProLeu 367
Db 3166 GGTGGGGACCTGACACTGAGGCTGGAGCCCTCTGAAGAGAGGCCCCAGGCTTCCACTG 3225
QY 368 AlaProSerGluGlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAla 387
Db 3226 GCACCTCCGAAGGGGCTGGCTCCGATGTATTGTATGTTGACCTGGGAATGGGGCAGCC 3285
QY 388 LysGlyLeuGlnSerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAsp 407
Db 3286 AAGGGGCTGCAAGGCTCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAGAGAC 3345
QY 408 ProThrValProLeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerPro 427
Db 3346 CCCACAGTACCCCTGCTCCCTCTGAGACTGATGCTACGTTGCCCTGACCTGCAGCCCC 3405
QY 428 GlnProGluTyrValAsnGlnProAspValArgProGlnProProSerProArgGluGly 447
Db 3406 CAGCCTGAATATGTGACACAGCATGTGCGCCCCAGCCCCCTTCGCCCGAGAGGGC 3465
QY 448 ProLeuProAlaAlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerPro 467
Db 3466 CCTGCTGCTGCTGCCGACCTGCTGTGTCACCTGTGAAGGGCCAGACTCTCTCCCCA 3525
QY 468 GlyLysAsnGlyValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGlu 487
Db 3526 GGGAAGATGGGGTGTCTAAAGACTTTTGGCTTTGGGGTCCGTGAGAACCCCGAG 3585
QY 488 TyrLeuThrProGlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerPro 507
Db 3586 TACTTGACACCCAGAGGAGAGCTGCCCTCAGCCCACTCTGCTTCAGCCCA 3645
QY 508 AlaPheAspAsnLeuTyrTyrTyrAspGlnAspProProGluArgGlyAlaProProSer 527
Db 3646 GCGTTCGACAACTCTTATTACTGGGACAGACCCAGAGCGGGGGCTCCACCCAGC 3705
QY 528 ThrPheLysGlyThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValPro 546
Db 3706 ACCTTCAAAGGACACCTACGGCAGAGAACCCAGAGTACTGGTCTGAGAGTGCCA 3762

Search completed: May 8, 2003, 14:16:48
Job time : 4411 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:40:24 ; Search time 25.3846 Seconds
(without alignments)
2178.074 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLFLF.....EPVOEGAPPPRAAHNNHHH 690

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3739	100.0	690	10 US-09-821-883-2	Sequence 2, Appl1
2	3473.5	92.9	697	10 US-09-821-883-4	Sequence 4, Appl1
3	2972	79.5	555	10 US-09-821-883-1	Sequence 1, Appl1
4	2957.5	79.1	564	10 US-09-821-883-3	Sequence 3, Appl1
5	2573.5	68.8	919	9 US-09-854-356-6	Sequence 6, Appl1
6	2410.5	64.5	1255	9 US-09-769-508-2	Sequence 2, Appl1
7	2410.5	64.5	1255	10 US-09-811-123-9	Sequence 9, Appl1
8	2410.5	64.5	1255	10 US-09-811-115-3	Sequence 1, Appl1
9	2405.5	64.3	1255	9 US-09-854-356-1	Sequence 1, Appl1
10	2405.5	64.3	1255	9 US-09-930-125-2	Sequence 2, Appl1
11	2405.5	64.3	1255	9 US-09-441-411-6	Sequence 5, Appl1
12	2188.5	58.5	479	10 US-09-821-883-5	Sequence 5, Appl1
13	1961	52.4	1256	9 US-09-854-356-14	Sequence 14, Appl1
14	1950.5	52.2	1260	9 US-09-870-759-118	Sequence 118, Appl1
15	1948	52.1	1256	9 US-09-854-356-2	Sequence 2, Appl1
16	1612	43.1	653	10 US-09-821-161-1	Sequence 1, Appl1
17	1612	43.1	653	9 US-09-854-356-3	Sequence 3, Appl1
18	1612	43.1	712	9 US-09-854-356-7	Sequence 7, Appl1
19	1587	42.4	289	10 US-09-821-883-23	Sequence 23, Appl1

20	1367.5	36.6	654	9 US-09-854-356-8	Sequence 8, Appl1
21	1183	31.6	566	9 US-09-854-356-4	Sequence 4, Appl1
22	1183	31.6	583	9 US-09-930-125-9	Sequence 9, Appl1
23	1183	31.6	589	9 US-09-930-125-8	Sequence 8, Appl1
24	1183	31.6	589	9 US-09-930-125-10	Sequence 10, Appl1
25	1183	31.6	600	9 US-09-930-125-11	Sequence 11, Appl1
26	1182	31.6	217	10 US-09-821-883-25	Sequence 25, Appl1
27	1182	31.6	397	10 US-09-821-883-27	Sequence 27, Appl1
28	1182	31.6	1179	10 US-09-821-883-29	Sequence 29, Appl1
29	893	23.9	191	9 US-09-441-411-9	Sequence 9, Appl1
30	721	19.3	657	9 US-10-172-620-18	Sequence 18, Appl1
31	721	19.3	1210	10 US-09-725-433-2	Sequence 2, Appl1
32	721	19.3	1308	10 US-09-940-101-2	Sequence 2, Appl1
33	717	19.2	293	9 US-10-102-806-583	Sequence 583, App
34	713.5	19.1	615	10 US-09-940-101-4	Sequence 4, Appl1
35	704.5	18.8	478	10 US-09-867-521-2	Sequence 2, Appl1
36	697	18.6	1342	9 US-10-172-620-16	Sequence 16, Appl1
37	684	18.3	610	10 US-09-783-708-1	Sequence 1, Appl1
38	674	18.0	144	10 US-09-923-246-114	Sequence 114, App
39	673	18.0	127	9 US-09-792-793A-15	Sequence 15, Appl1
40	673	18.0	127	10 US-09-821-883-18	Sequence 18, Appl1
41	673	18.0	127	10 US-09-800-016-1	Sequence 1, Appl1
42	648.5	17.3	135	10 US-09-925-301-1232	Sequence 1232, Ap
43	441	11.8	127	10 US-09-821-883-20	Sequence 20, Appl1
44	373.5	10.0	219	10 US-09-847-185-2	Sequence 2, Appl1
45	255	6.8	120	9 US-10-172-620-17	Sequence 17, Appl1

ALIGNMENTS

RESULT 1
US-09-821-883-2
Sequence 2, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Reiner
APPLICANT: Vidovics, Tamir
APPLICANT: Graddis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821, 883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 690
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500-HCM-CSF construct
US-09-821-883-2
Query Match 100.0%; Score 3739; DB 10; Length 690;
Best Local Similarity 100.0%; Pred. No. 4e-212;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MRAAPLLARAASLSGLFLF...FWLDRSVLAKELARGAASVOTGTDMLRLPASPT 60
DB 1 MRAAPLLARAASLSGLFLF...FWLDRSVLAKELARGAASVOTGTDMLRLPASPT 60
OY 61 HMDMRLHYOGCGVVOGNELETLPTNASTSLFDIOIEYGVYVLAHNOVROPVLPORLRI 120
DB 61 HMDMRLHYOGCGVVOGNELETLPTNASTSLFDIOIEYGVYVLAHNOVROPVLPORLRI 120
OY 121 VAGTOLFEDNVALAVLDNGDPLNNTPTVGTASPGGLRELOJRSLEIKGGVLIORNPOL 180
DB 121 VAGTOLFEDNVALAVLDNGDPLNNTPTVGTASPGGLRELOJRSLEIKGGVLIORNPOL 180
OY 181 CYODTILMKDIFHKNNOLATLTLIDITNRSRACHPCSPMKSGKRCNCESSDSCSLTRTVCA 240

```

Db 181 CYDDTILMKDIFKHNKNOALATLIDTNRSRACHCSPKCSKRCWGESSEDCOSLTRVCA 240
QY 241 GGCARCKGRLPTDCHCBOCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
Db 241 GGCARCKGRLPTDCHCBOCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
QY 301 MPNPEGRTFGASCVTACPNYLSTDVSGAGGVNHHRRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNPEGRTFGASCVTACPNYLSTDVSGAGGVNHHRRSSSTRSGGDLTLGLEPSEE 360
QY 361 EAPRSPLABEGASDVFDODLGMGAAGLQSLPTHDPSPLOQRYSEPTVPLPSEIDGYV 420
Db 361 EAPRSPLABEGASDVFDODLGMGAAGLQSLPTHDPSPLOQRYSEPTVPLPSEIDGYV 420
QY 421 APLTCSQPEYVNOQDVRRPQPSRREGPLPAPRAGATLERAKTILSPGKGVKDVPAFG 480
Db 421 APLTCSQPEYVNOQDVRRPQPSRREGPLPAPRAGATLERAKTILSPGKGVKDVPAFG 480
QY 481 GAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLYYWDDPPERGAPSPSTFKGTPAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLYYWDDPPERGAPSPSTFKGTPAENPEY 540
QY 541 LGIDVPAARASPSSTOQWENHVAIOEARLLNLSRDTAAEMNETVEYISEMFDLOEP 600
Db 541 LGIDVPAARASPSSTOQWENHVAIOEARLLNLSRDTAAEMNETVEYISEMFDLOEP 600
QY 601 TCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITFESFKENLKD 660
Db 601 TCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITFESFKENLKD 660
QY 661 FLVYIPDCMEPVQEGAPPPAAAHNNHH 690
Db 661 FLVYIPDCMEPVQEGAPPPAAAHNNHH 690

RESULT 2
US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Gradalis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct
US-09-821-883-4

```

```

Query Match 92.9%; Score 3473.5; DB 10; Length 697;
Best Local Similarity 91.7%; Pred. No. 1.6e-196;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;

```

```

QY 1 MRAAPLLAARASLSGFLFLEFMDRSYLAKELARGASTOVCTGDMKLRPASPER 60
Db 1 MRAAPLLAARASLSGFLFLEFMDRSYLAKELARGASTOVCTGDMKLRPASPER 60
QY 61 HLDMLHLYOGCGVGNLELTPTNASLSFODIOEVGYVLIHNOVROVPLQRLRI 120
Db 61 HLDMLHLYOGCGVGNLELTPTNASLSFODIOEVGYVLIHNOVROVPLQRLRI 120

```

```

QY 121 VRGTQLEEDYVALAVLNDGNPLNNTTPTVTGASPGGLRELQRLSTELKGVLIQRRPOL 180
Db 121 VRGTQLEEDYVALAVLNDGNPLNNTTPTVTGASPGGLRELQRLSTELKGVLIQRRPOL 180
QY 181 CYDDTILMKDIFKHNKNOALATLIDTNRSRACHCSPKCSKRCWGESSEDCOSLTRVCA 240
Db 181 CYDDTILMKDIFKHNKNOALATLIDTNRSRACHCSPKCSKRCWGESSEDCOSLTRVCA 240
QY 241 GGCARCKGRLPTDCHCBOCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
Db 241 GGCARCKGRLPTDCHCBOCAAGCTGPKHSDCLCLHFNHSGICELHCPALVTYNTDFES 300
QY 301 MPNPEGRTFGASCVTACPNYLSTDVGS-----GAGGVNHHRRSSSTRSGGDL 351
Db 301 MPNPEGRTFGASCVTACPNYLSTDVGSASINFEKLAGGVNHHRRSSSTRSGGDL 360
QY 352 TLGLEPSEEEAPRSPLABSGASDVFDODLGMGAAGLQSLPTHDPSPLOQRYSEPTV 411
Db 352 TLGLEPSEEEAPRSPLABSGASDVFDODLGMGAAGLQSLPTHDPSPLOQRYSEPTV 420
QY 412 LPSETDGYVAPLTCSPQPEYVNOQDVRRPQPSRREGPLPAPRAGATLERAKTILSPGKNG 471
Db 421 LPSETDGYVAPLTCSPQPEYVNOQDVRRPQPSRREGPLPAPRAGATLERAKTILSPGKNG 480
QY 472 VKQDVEAFGAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLYYWDDPPERGAPSPSTFKG 531
Db 481 VKQDVEAFGAVENPEYLTPOGGAAPQHPHPAPFSPAFDNLYYWDDPPERGAPSPSTFKG 540
QY 532 TPNAENPEYLGIDVPAARASPSSTOQWENHVAIOEARLLNLSRDTAAEMNETVEYI 591
Db 541 TPNAENPEYLGIDVPAARASPSSTOQWENHVAIOEARLLNLSRDTAAEMNETVEYI 600
QY 592 SEMFDLOEPTCLQTRLELYKQGLRSLTKLGPLTMMASHYKONCPPTETSCATOIITF 651
Db 601 SNESIQRPCTQVQTRKLYKQGLRSLTKLGNLTKLNGALTMASHYKONCPPTETSCATOIITF 660
QY 652 ESFKENLKDPLVYIPDCMEPVQEGAPPPAAAHNNHH 690
Db 661 EDFIKNLKGLFDIPEDCMWRVQKGAAPP--AHNNHH 697

```

```

RESULT 3
US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Gradalis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

```

```

Query Match 79.5%; Score 2972; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 3.6e-167;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRAAPLLAARASLSGFLFLEFMDRSYLAKELARGASTOVCTGDMKLRPASPER 60
Db 1 MRAAPLLAARASLSGFLFLEFMDRSYLAKELARGASTOVCTGDMKLRPASPER 60

```

```

QY 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLDIOEVQGYVLIANQVROVPLQRLRI 120
DB 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLDIOEVQGYVLIANQVROVPLQRLRI 120
QY 121 VRGTQLEFDNVALAVLNDGDP LNNTTPTVTGASPGGLRELQRLSLTEILKGVLLQRLNOL 180
DB 121 VRGTQLEFDNVALAVLNDGDP LNNTTPTVTGASPGGLRELQRLSLTEILKGVLLQRLNOL 180
QY 181 CYODTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESESDCOSLTRVCA 240
DB 181 CYODTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESESDCOSLTRVCA 240
QY 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300
DB 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300
QY 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVYHHRHRSSTSGGDLTLGLEPSEE 360
DB 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVYHHRHRSSTSGGDLTLGLEPSEE 360
QY 361 EAPRSPLAPSRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVLPSETDGYV 420
DB 361 EAPRSPLAPSRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVLPSETDGYV 420
QY 421 APLTCSPOPEYVNPDPVRPQPPSPREGPLPAARPAAGATLERAKTILSPKNGVVKDVFAG 480
DB 421 APLTCSPOPEYVNPDPVRPQPPSPREGPLPAARPAAGATLERAKTILSPKNGVVKDVFAG 480
QY 481 GAVENPEYLTPOGGAAPQHPHPAFSPAEDNLYWDDPBERGAPSTFKGTPAENDEY 540
DB 481 GAVENPEYLTPOGGAAPQHPHPAFSPAEDNLYWDDPBERGAPSTFKGTPAENDEY 540
QY 541 LGLDVPAAP 549
DB 541 LGLDVPAAP 549

```

RESULT 4

```

US-09-821-883-3
: Sequence 3, Application US/09821883
: Patent No. US20020061310A1
: GENERAL INFORMATION:
: APPLICANT: Laus, Retner
: APPLICANT: Givodis, Thamar
: APPLICANT: Vidovic, Thomas
: TITLE OF INVENTION: Compositions and Methods for Dendritic
: FILE REFERENCE: 7636-0022.30
: CURRENT APPLICATION NUMBER: US/09/821,883
: PRIOR FILING DATE: 2001-03-30
: PRIOR APPLICATION NUMBER: US 60/193,504
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 564
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: HER500* construct
US-09-821-883-3

```

Query Match Best Local Similarity 98.4%; Score 2957.5; DB 10; Length 564;

Matches 549; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

```

QY 1 MRAAPLLARASISLGLFLFLFWLDRSVAKELARGASTOVCTGDMKRLRPASPE 60
DB 1 MRAAPLLARASISLGLFLFLFWLDRSVAKELARGASTOVCTGDMKRLRPASPE 60
QY 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLDIOEVQGYVLIANQVROVPLQRLRI 120
DB 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLDIOEVQGYVLIANQVROVPLQRLRI 120

```

```

DB 61 HLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLDIOEVQGYVLIANQVROVPLQRLRI 120
QY 121 VRGTQLEFDNVALAVLNDGDP LNNTTPTVTGASPGGLRELQRLSLTEILKGVLLQRLNOL 180
DB 121 VRGTQLEFDNVALAVLNDGDP LNNTTPTVTGASPGGLRELQRLSLTEILKGVLLQRLNOL 180
QY 181 CYODTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESESDCOSLTRVCA 240
DB 181 CYODTILMKDIFHNKNQALATLIDTNRSRACHPCSPMKGSRGCESESDCOSLTRVCA 240
QY 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300
DB 241 GGCARCKGRLPTDCHECCAGCTGPKHSDCLACLFHNHSGICELHCPALTYNTDPTES 300
QY 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVYHHRHRSSTSGGDLTLGLEPSEE 360
DB 301 MPNEGRTFGASCVTACPYNYLSTDVSGAGVYHHRHRSSTSGGDLTLGLEPSEE 360
QY 361 EAPRSPLAPSRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVLPSETDGYV 420
DB 361 EAPRSPLAPSRGASDVFDGDLGMAKGLQSLPTHDPSPLORYSEDPVLPSETDGYV 420
QY 421 APLTCSPOPEYVNPDPVRPQPPSPREGPLPAARPAAGATLERAKTILSPKNG 471
DB 421 APLTCSPOPEYVNPDPVRPQPPSPREGPLPAARPAAGATLERAKTILSPKNG 471
QY 471 VVKDVFAGVAVENPEYLTPOGGAAPQHPHPAFSPAEDNLYWDDPBERGAPSTFKG 531
DB 471 VVKDVFAGVAVENPEYLTPOGGAAPQHPHPAFSPAEDNLYWDDPBERGAPSTFKG 531
QY 532 TPTAENPEYLGIDVPAAP 549
DB 532 TPTAENPEYLGIDVPAAP 549
QY 541 LGLDVPAAP 558
DB 541 LGLDVPAAP 558

```

RESULT 5

```

US-09-854-356-6
: Sequence 6, Application US/09854356
: Patent No. US2002017567A1
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Gheysen, Dirk
: APPLICANT: Corixa Corporation
: APPLICANT: SmithKline Beecham Biologicals S. A.
: TITLE OF INVENTION: HER-2/neu Fusion Proteins
: FILE REFERENCE: 014058-009810PC
: CURRENT APPLICATION NUMBER: US/09/854,356
: PRIOR FILING DATE: 2001-05-09
: PRIOR APPLICATION NUMBER: US 09/493,480
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: US 60/117,976
: PRIOR FILING DATE: 1999-01-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 919
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: fusion protein
: OTHER INFORMATION: of ECD and PD of human HER-2/neu
US-09-854-356-6

```

Query Match Best Local Similarity 68.8%; Score 2573.5; DB 9; Length 919;

Matches 509; Conservative 0; Mismatches 3; Indels 391; Gaps 1;

```

QY 35 LARGAASOVCTGDMKRLRPASPEHLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLQ 94
DB 16 LARGAASOVCTGDMKRLRPASPEHLDMLRHLHYGCGVYVGNLELTLYLPTNASTLSFLQ 94
QY 95 DIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLNDGDP LNNTTPTVTGASPG 154
DB 95 DIOEVQGYVLIANQVROVPLQRLRIYRGTOLEFDNVALAVLNDGDP LNNTTPTVTGASPG 154

```

```

Db 76 DIOEVGYVLIHNOVQVPLQRLRVGTQLEEDNVALVLDNGDPLNNTPTVTGASPG 135
QY 155 GLRELQRLSTLEILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 214
Db 136 GLRELQRLSTLEILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 195
QY 215 SPMCKSRGWESSEDCQSLRTVACAGCARCKGRLPTDCHEQCAAGCTGPRHSDCLAC 274
Db 196 SPMCKSRGWESSEDCQSLRTVACAGCARCKGRLPTDCHEQCAAGCTGPRHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDFFESMPNPEGRTTGASCVTACPYNYLSTDVGSCTLYC 329
Db 256 LHFHNSGICELHCPALVTYNTDFFESMPNPEGRTTGASCVTACPYNYLSTDVGSCTLYC 315
QY 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREAVRATSANIOEFAGCKKIEGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLQBPOLQVETLEETGYLYISAMPDLPDLVSFQNLQYIGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTLOGIGISWGLRLSRELGSGLALIHNTNHLCTVHTVPMQDLFRNPQALH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGLAGHQLCARGHCWGPRTQCVCNSQFLRGQCEVBECAVLQGLPREV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVACRPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEBACQCPINCHTSCVDLDKCPARQASPLTSQNEGLAPASPLDSTYRSLLED 675
QY 330 ----- 329
Db 676 DMGDLVDADEYLVPOQGFPCRPDAPAGMHNHNRSSSTRSGGDLTGLPSESEAP 735
QY 364 RSLPLASEGASVDFDGLMGAAKGLQSLPTDPSPLQYSEDPTVYRLPSEDTGYAPL 423
Db 736 RSLPLASEGASVDFDGLMGAAKGLQSLPTDPSPLQYSEDPTVYRLPSEDTGYAPL 795
QY 424 TCSPOEYVNOQVPRQPSRPGPLPAARPAATLERAKTSLPGKNVYKDVFAFGAV 483
Db 796 TCSPOEYVNOQVPRQPSRPGPLPAARPAATLERAKTSLPGKNVYKDVFAFGAV 855
QY 484 ENPEYLTPOGGAAPQHPAPAFSPADNLYYMDODPERGAPSTFGKPTAENPEYLG 543
Db 856 ENPEYLTPOGGAAPQHPAPAFSPADNLYYMDODPERGAPSTFGKPTAENPEYLG 915
QY 544 DVP 546
Db 916 DVP 918

```

RESULT 6
US-09-769-508-2

Sequence 2, Application US/09769508
Patent No. US20020155527A1
GENERAL INFORMATION:
APPLICANT: STUART, SUSAN G.
APPLICANT: MONAHAN, JOHN J.
APPLICANT: LANGTON, BEATRICE CLAUDIA
APPLICANT: HANCOCK, MIRIAM E.C.
APPLICANT: CHAO, LORINE A.
APPLICANT: BLUFORD, PETER
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
FILE REFERENCE: BBIO-111-C1
CURRENT APPLICATION NUMBER: US/09/769,508

```

; CURRENT FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-769-508-2

Query Match      64.5%; Score 2410.5; DB 9; Length 1255;
Best Local Similarity 41.2%; Pred. No. 9.1e-134;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGASTOYCTGTDMLRLPASPEHLDMLRHLQYQVYVGNLELYLPTNASLSFQ 94
Db 16 LPGAASTOYCTGTDMLRLPASPEHLDMLRHLQYQVYVGNLELYLPTNASLSFQ 75
QY 95 DIOEVGYVLIHNOVQVPLQRLRVGTQLEEDNVALVLDNGDPLNNTPTVTGASPG 154
Db 76 DIOEVGYVLIHNOVQVPLQRLRVGTQLEEDNVALVLDNGDPLNNTPTVTGASPG 135
QY 155 GLRELQRLSTLEILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 214
Db 136 GLRELQRLSTLEILKGVLIQIRNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 195
QY 215 SPMCKSRGWESSEDCQSLRTVACAGCARCKGRLPTDCHEQCAAGCTGPRHSDCLAC 274
Db 196 SPMCKSRGWESSEDCQSLRTVACAGCARCKGRLPTDCHEQCAAGCTGPRHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDFFESMPNPEGRTTGASCVTACPYNYLSTDVGSCTLYC 329
Db 256 LHFHNSGICELHCPALVTYNTDFFESMPNPEGRTTGASCVTACPYNYLSTDVGSCTLYC 315
QY 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSPCARVCYGLGMEHLREAVRATSANIOEFAGCKKIEGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLQBPOLQVETLEETGYLYISAMPDLPDLVSFQNLQYIGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTLOGIGISWGLRLSRELGSGLALIHNTNHLCTVHTVPMQDLFRNPQALH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGLAGHQLCARRALLSGPTQCVCNSQFLRGQCEVBECAVLQGLPREV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECOPONGSVTCFGEADQVACAHYKDPFCVACRPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEBACQCPINCHTSCVDLDKCPARQASPLTSIVSAVGLLVVVGVEGILT 675
QY 330 ----- 329
Db 676 KRROQIRKTYMRLLQETELVEPLTPSGAMPNOAMRIKETEELKRVKVLGSGAGCTV 735
QY 330 ----- 329
Db 736 KGIWPDGENVKIPVAKVIRENTSPRANKELIDEAVVAGVSPYVSRLLGICLSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGLLDHVRNRRGLSODLLNMCQIANGMSYLEDRVLVHRDLAARNLVKS 855
QY 330 ----- 329
Db 856 PNHVKTIDFGLALLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSYGVYVWEL 915

```

QY 330 ----- 329
Db 916 MTFGAKPYDGIIPAREIPDLLEKGERLQPPICITIDVYIMVCMIMIDSECRPRELVE 975
QY 330 ----- 329
Db 976 FSRMARDPQRFVITONEDLGPASPLDSTFYRSLLLEDDMDGLVDAEEXLVPOGFFCDDP 1035
QY 330 --GAGGWHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 387
Db 1036 APGAGWHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 1095
QY 388 KGLQSLPHTDPSLQRYSEDPVLPSETDGYVAPLTCSPQPEYVNOQDVAPRPPSPREG 447
Db 1096 KGLQSLPHTDPSLQRYSEDPVLPSETDGYVAPLTCSPQPEYVNOQDVAPRPPSPREG 1155
QY 448 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEVLTPOGGAAPRPHPPAPSP 507
Db 1156 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEVLTPOGGAAPRPHPPAPSP 1215
QY 508 AFDNLXYWDODPPERGAPSTFKGTPTAENPEYLGIDVP 546
Db 1216 AFDNLXYWDODPPERGAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 7

US-09-811-123-9

: Sequence 9, Application US/09811123
: Patent No. US2002001587A1
: GENERAL INFORMATION:
: APPLICANT: Sharon Erickson
: APPLICANT: Ralph Schwall
: APPLICANT: Mark Sliwowski
: TITLE OF INVENTION: METHODS OF TREATMENT USING ANTI-EBDB
: FILE REFERENCE: GENENT.073A2
: CURRENT APPLICATION NUMBER: US/09/811.123
: CURRENT FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/238,327
: PRIOR FILING DATE: 2000-10-05
: PRIOR APPLICATION NUMBER: 09/602,530
: PRIOR FILING DATE: 2000-06-23
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 9
: LENGTH: 1255
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-811-123-9

Query Match 64.5% Score 2410.5; DB 10; Length 1255;
Best Local Similarity 41.2%; Pred. No. 9.1e-134;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;
QY 35 LARGAASVCTGTDMKRLPASPETHLMDRLHYCQGVVQGNLELTYLPTNASTLFLQ 94
Db 16 LPPGAASVCTGTDMKRLPASPETHLMDRLHYCQGVVQGNLELTYLPTNASTLFLQ 75
QY 95 DIOEVQGVLLAHNOVQRYPLQRLRYNGTOLFEDNVALAVLDNGDPLNNTTPTVGASPG 154
Db 76 DIOEVQGVLLAHNOVQRYPLQRLRYNGTOLFEDNVALAVLDNGDPLNNTTPTVGASPG 135
QY 155 GLRELOLRLSLTEILKGLVLIORNPOLCYODTILMKDIFHKNNOLALTLIDINRSRACHPG 214
Db 136 GLRELOLRLSLTEILKGLVLIORNPOLCYODTILMKDIFHKNNOLALTLIDINRSRACHPG 195
QY 215 SPMCKSGRCWGESSEDCSLRTYVACAGCARKGRLPTDCHEQCAACCTGPKHSDCLAC 274
Db 196 SPMCKSGRCWGESSEDCSLRTYVACAGCARKGRLPTDCHEQCAACCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVYVNTDTEFSMNPREGRTTFGASCTYACRYNLTSDVGS----- 329
Db 256 LHFHNSGICELHCPALVYVNTDTEFSMNPREGRTTFGASCTYACRYNLTSDVGSCTLVG 315

QY 330 ----- 329
Db 316 PLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIQEFAGCKRTFGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGPASNTAPLQPEQLQVFEYLEITGYLTISAMPDPLPDLVSFQNLQVIRGRI 435
QY 330 ----- 329
Db 436 LHNAGYSLTLOGLGI SWGLNLSRELGSGLALIHNNTHLCFVHTVPMQDLFRNPQALLH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHOLCARGHCMRGPTQVCYNCSQFLRGQCEVECRVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPEQOPONGSVTCGPEADQCVACAHYKDPFCVARKPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 PPDEGACQPCPINCTHSCVDLDDKCPAEQASPLSYSAVVGILLVVLGVVFGILI 675
QY 330 ----- 329
Db 676 KRROOKIRKTYMRLOETELVEPLTPSGAMPNOAOMRILKETELRKXKVLGSCAFGIVY 735
QY 330 ----- 329
Db 736 KGIWIPQENYKIPVAKVIRENTSPKANKELIDEAVMAGVSPYVRLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTLQMPYCLLDHYRENRGLGSODLLMCMQIAKMSYLEDVRLVHRDLAARVVLKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGCKVPIKMALESILRRRTFHOSDVSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIIPAREIPDLLEKGERLQPPICITIDVYIMVCMIMIDSECRPRELVE 975
QY 330 ----- 329
Db 976 FSRMARDPQRFVITONEDLGPASPLDSTFYRSLLLEDDMDGLVDAEEXLVPOGFFCDDP 1035
QY 330 --GAGGWHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 387
Db 1036 APGAGWHHRHRSSTRSGGDLTLGLEPSEEARPSPLAPSBGAGSDVFDGDLGMAA 1095
QY 388 KGLQSLPHTDPSLQRYSEDPVLPSETDGYVAPLTCSPQPEYVNOQDVAPRPPSPREG 447
Db 1096 KGLQSLPHTDPSLQRYSEDPVLPSETDGYVAPLTCSPQPEYVNOQDVAPRPPSPREG 1155
QY 448 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEVLTPOGGAAPRPHPPAPSP 507
Db 1156 PLPARAPGATLERAKTILSPKNGVYKDVFAFGAVENPEVLTPOGGAAPRPHPPAPSP 1215
QY 508 AFDNLXYWDODPPERGAPSTFKGTPTAENPEYLGIDVP 546
Db 1216 AFDNLXYWDODPPERGAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 8

US-09-811-115-3

: Sequence 3, Application US/09811115
: Patent No. US20020035736A1
: GENERAL INFORMATION:
: APPLICANT: Erickson, Sharon
: APPLICANT: Schwall, Ralph

APPLICANT: King, Kathleen
TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
FILE REFERENCE: GENENT. 034A
CURRENT APPLICATION NUMBER: US/09/811,115
CURRENT FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/189,844
PRIOR FILING DATE: 2000-03-16
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1255
TYPE: PRT
ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 64.5%; Score 2410.5; DB 10; Length 1255;
Best Local Similarity 41.1%; Pred. No. 9.1e-134;

Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGASTOYCTGTMKRLPASPETHLDMRLHYOGCQVYVGNLELTYLPTNASLSFLQ 94
D 16 LPGAASTOYCTGTMKRLPASPETHLDMRLHYOGCQVYVGNLELTYLPTNASLSFLQ 75
QY 95 DIOGNGYVLIANNQVYVLOLRLRYGTQLEEDNATALVLDNGPPLNTTPTTGASPG 134
D 76 DIOGNGYVLIANNQVYVLOLRLRYGTQLEEDNATALVLDNGPPLNTTPTTGASPG 135
QY 155 GLRELQRLSTLTLKGVLLQIQRNPOLCYODTIIMKDIFFHNNOLATLIDITNRSRACHPC 214
D 136 GLRELQRLSTLTLKGVLLQIQRNPOLCYODTIIMKDIFFHNNOLATLIDITNRSRACHPC 135
QY 215 SPKCKSRCKGSESEDCQSLTRTVYVCAAGCARGCGLPTDCCHQCAAGCTGPKHSDCLAC 274
D 196 SPKCKSRCKGSESEDCQSLTRTVYVCAAGCARGCGLPTDCCHQCAAGCTGPKHSDCLAC 255
QY 275 LPHNHSIGICELHCPALVTYVTFEESMPNPEGRTTGASCVTACPNYVSTDVGS----- 329
D 256 LPHNHSIGICELHCPALVTYVTFEESMPNPEGRTTGASCVTACPNYVSTDVGSCTLVC 315
QY 330 ----- 329
D 316 PLHNGEYTAEDGTQCEKCKPCARVYCYGLMEHLREAVRVSANTIOEFAGCKKIGSLA 375
QY 330 ----- 329
D 376 FLPESEFGDPASNTAPLOPEOLOVFETLEETGYLYISAMPDSLPLDSYFQNLQVIRGRI 435
QY 330 ----- 329
D 436 LHHGAVSLTLQGLISMLGRLSLRELGLALIHNTHTLFCFHTVPMQDLFRNPQALLH 495
QY 330 ----- 329
D 496 TANRPDECVSEGLACHQCLARGHCWGPPTQCVNCSQFLRGQECVECHVLOGLPREYV 555
QY 330 ----- 329
D 556 NARHCLPCHPECOPIONGSVTCFSPREADOCVACAHYKDPFCVACRPSGVKPLSYPMK 615
QY 330 ----- 329
D 616 FPBEGACOPCPINCTHSCVDLDDKCPABQKASPLSTIYSANVAILLVVLCVFGIL 675
QY 330 ----- 329
D 676 KRROKIRKTYMRRLQETELVEPLTPSGAMPNQAMRLKETELRKVKVLGSGAGFTYV 735
QY 330 ----- 329
D 736 KGIWIDGENVKIPVAKVLENTSPKANKEILDEAVYVAGVSPYVSRLLGICLTSTVO 795
QY 330 ----- 329

D 796 LVTQLMPYGLCLLDHVENRGRRLGSDLLNMCQIAKMSYLEYVLRVHDLAARNLVYS 855
QY 330 ----- 329
D 856 PNHVKTIDFGLARLDIDETEVHADGKVPKIMMALESILRRRTHQSDVHSGVYVWEL 915
QY 330 ----- 329
D 916 MTEGAKPYDIPAREIPDLLEKGERLPPQPICTIDVYIMVKGWIMIDSECRPRELYSE 975
QY 330 ----- 329
D 976 FSRMARDPQRFVVIQNEEDLCPASPLDSTFYRLSLEDMDKDLVDAEYLYPQGFECPPD 1035
QY 330 ----- 329
D 1036 APGAGGVHHRHSSSTRSGGDLTLGLSESEEPASPPLAPSEAGSDYFDDGLGMAA 1095
QY 388 KGLQSLPTHPSPLOQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVPPSPREG 447
D 1096 KGLQSLPTHPSPLOQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQDVPPSPREG 1155
QY 448 PLPAARAGATLERAKTLSGKNGVYVDFAFGAVENPEYLPPOGAAPQPPPAFSP 507
D 1156 PLPAARAGATLERAKTLSGKNGVYVDFAFGAVENPEYLPPOGAAPQPPPAFSP 1215
QY 508 AFNLYVWDODPPERGAPSTFKGTPTAENPEYLGIDVP 546
D 1216 AFNLYVWDODPPERGAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 9

US-09-854-356-1

Sequence 1, Application US/09854356

Patent No. US20020177567A1

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Gneysen, Dirk

APPLICANT: Corixa Corporation

APPLICANT: SmithKline Beecham Biologicals S. A.

TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC

CURRENT APPLICATION NUMBER: US/09/854,356

CURRENT FILING DATE: 2001-05-09

PRIOR APPLICATION NUMBER: US 09/493,480

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: US 60/117,976

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 26

SOFTWARE: Patentlin Ver. 2.1

SEQ ID NO 1

LENGTH: 1255

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: human HER-2/neu protein

NAME/KEY: DOMAIN

LOCATION: (1)..(653)

OTHER INFORMATION: extracellular domain (ECD)

NAME/KEY: DOMAIN

LOCATION: (676)..(1255)

OTHER INFORMATION: Intracellular domain (ICD)

NAME/KEY: DOMAIN

LOCATION: (990)..(1255)

OTHER INFORMATION: phosphorylation domain (PD)

NAME/KEY: DOMAIN

LOCATION: (990)..(1048)

OTHER INFORMATION: fragment of the phosphorylation domain, preferred

US-09-854-356-1

Query Match 64.3%; Score 2405.5; DB 9; Length 1255;
Best Local Similarity 41.1%; Pred. No. 1.8e-133;

[illegible]

```

0Y      388      KGLSLPHDPSLQRYSEDPVPLPSEFDGVAALVLTCSPOREYNQOPVRRQPSREG  447
Db      1096     KGLSLPHDPSLQRYSEDPVPLPSETDGVAAVLTCSPOREYNQOPVRRQPSREG  1155
0Y      448     PLPARPAGATLERKKTLSPGKNVGVKDVFAFGAVENPEYLTTPGGAAPORHPAPASP  507
Db      1156     PLPARPAGATLERKKTLSPGKNVGVKDVFAFGAVENPEYLTTPGGAAPORHPAPASP  1215
0Y      508     AFDNLVYWDQDPPERGAPSPSTFKGPTPLANPEYLGLDVP  546
Db      1216     AFDNLVYWDQDPPERGAPSPSTFKGPTPLANPEYLGLDVP  1254

RESULT 10
US-09-930-125-2
; Sequence 2, Application US/09930125
; Publication No. US20020193329A1
; GENERAL INFORMATION:
; APPLICANT: Hand-Zimmerman, Susan
; APPLICANT: Cheever, Martin A.
; APPLICANT: Foy, Teresa M.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Veevick, Thomas S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
; FILE REFERENCE: 210121.544
; CURRENT APPLICATION NUMBER: US/09/930.125
; CURRENT FILING DATE: 2001-08-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-930-125-2

Query Match      64.3%; Score 2405.5; DB 9; Length 1255;
Best Local Similarity 41.1%; Pred. No. 1.8e-133;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

0Y      35      LARGASTOYVCTGMKRLPASPETHLDMRLHYOGCOVQGNLELTPLTNASLSFLQ  94
Db      16      LPPGASTOYVCTGDTMKRLPASPETHLDMRLHYOGCOVQGNLELTPLTNASLSFLQ  75
0Y      95      DIOEVQGVYLLAHNOVROYPLQRLRIYRGTOLEFDNYALAVLDNCGDPLNNTPVYGASFG  154
Db      76      DIOEVQGVYLLAHNOVROYPLQRLRIYRGTOLEFDNYALAVLDNCGDPLNNTPVYGASFG  135
0Y      155     GLERLQLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNQALVTLIDTNRSRACHC  214
Db      136     GLERLQLSLTEILKGVLIQRPOLCYQDTILMKDIFHKNNQALVTLIDTNRSRACHC  195
0Y      215     SPMCGSRMGWSSSDCCSLRTTYVAGGACARCKGFLPTDCCHEGCAAGTGKHSDDLAC  274
Db      196     SPMCGSRMGWSSSDCCSLRTTYVAGGACARCKGFLPTDCCHEGCAAGTGKHSDDLAC  255
0Y      275     LHFNHSICELCPALVYNTDTFESMPNPEGRTYFFGASCYACAPVNYLSTDVGS-----  329
Db      256     LHFNHSICELCPALVYNTDTFESMPNPEGRTYFFGASCYACAPVNYLSTDVGSCTLYC  315
0Y      330     -----  329
Db      316     PLHNQEVTAEDGTQRCCKSKPCARVCGLGMEHLREVRAYTSANIQEPAGCKIIFGSLA  375
0Y      330     -----  329
Db      376     FLPESEFDDPASNTAPLQPEQLQVETLEETIGVLYISAMPDLPDLVFQNIQVIRGRI  435
0Y      330     -----  329

```

```

Db 436 LHNAGVSLTGLGISMGLRSRLRELGSLALIHNTHLCEVHTVPWDLFRNPQALH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQCLARGHCMPGPTQCVCNCSQFLRGCEVECHVLOGLPREY 555
QY 330 ----- 329
Db 556 NARHCLPCHPECCOPONGSVTCFGEPEADOCVACAHRKDPFCVARCPSGVKPDLSYMPWK 615
QY 330 ----- 329
Db 616 FPDEBACQCPINCTHSCVDLDDKCCPABQRASPLTISIISAVGILLVVLGVGEGILI 675
QY 330 ----- 329
Db 676 KRQOKIRKTYMRRLQETELVEPLTPSGAMPQAOQMRILKETELRKVKVLSGARFTY 735
QY 330 ----- 329
Db 736 KGIWPDGENVKIPVAIKVLENTSPRANKELIDEAVYVAGSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCGLDHRVBNRGLSODLLMKMOIAKMSYLEVRLVHRDLAARNVLYKS 855
QY 330 ----- 329
Db 856 PNHVKTIDFGLARLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSGYVWEL 915
QY 330 ----- 329
Db 916 MTEGAPYDGIPIAREIPDLLEKGERLPPEICTIDVYIMVKCMTIDSECRPRFRELVE 975
QY 330 ----- 329
Db 976 FSRMADDPQRFVYIQNEDGLPASPDLSTFKSLLEDDMDGLVDAEYLVPOGFCPCPD 1035
QY 330 ----- 329
Db 1036 APGAGMHHRRSSSTRSGGDLTLGLEPSEBAPRSLAPSEGASDVFDGLGMA 1095
QY 330 ----- 329
Db 1096 KGIQSLPTDHPSPLOKSEDPYVPLPSETDGYAPLTCSPQREYVNOQDVRPQPSREG 1155
QY 448 PLFAPAPAGATLERAKTLSPGKNGVYKDVAFGAVENPEYLPQGAAPQHPAPFSP 507
Db 1156 PLFAPAPAGATLERAKTLSPGKNGVYKDVAFGAVENPEYLPQGAAPQHPAPFSP 1215
QY 508 AFDNLYYWDODPPERGAPSTFKGPTAENPEYLGLDVP 546
Db 1216 AFDNLYYWDODPPERGAPSTFKGPTAENPEYLGLDVP 1254

```

RESULT 11
US-09-441-411-6

```

: Sequence 6, Application US/09441411
: Publication NO. US20030008342A1
: GENERAL INFORMATION:
: APPLICANT: Scholler, Nathalie B.
: APPLICANT: Disis, Mary L.
: APPLICANT: Hellstrom, Ingegerd
: APPLICANT: Hellstrom, Karl Erik
: TITLE OF INVENTION: SURFACE RECEPTOR ANTIGEN VACCINES
: FILE REFERENCE: 730033.409
: CURRENT APPLICATION NUMBER: US/09/441,411
: CURRENT FILING DATE: 1999-11-16
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 1255
: TYPE: PRF

```

```

: ORGANISM: Homo sapiens
: US-09-441-411-6
Query Match 64.3% Score 2405.5; DB 9: Length 1255;
Best Local Similarity 41.1%; Pred. No. 1.8e-133;
Matches 509; Conserved 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGASTQVCTGDMKRLPASPETHLDMRLHLYOCQVVOGNLELTYLPNASSLFLQ 94
Db 16 LPPGAASVQVCTGDMKRLPASPETHLDMRLHLYOCQVVOGNLELTYLPNASSLFLQ 75
QY 95 DIOEVQGYVLIANNQVQVLORLRYRGQLEEDNIALALVNDGNPLNNTTYTGASPG 134
Db 76 DIOEVQGYVLIANNQVQVLORLRYRGQLEEDNIALALVNDGNPLNNTTYTGASPG 135
QY 155 GLRELOLRSLTEILKGVLLQORNPOLCYOPTILMKDIFHKNNOLATLIDTNSRACHPC 214
Db 136 GLRELOLRSLTEILKGVLLQORNPOLCYOPTILMKDIFHKNNOLATLIDTNSRACHPC 195
QY 215 SPMCKSRGWESSESDQSLTRTVACAGCARCKGPLETDCHEQCAAGCTGPRHSDCLAC 274
Db 196 SPMCKSRGWESSESDQSLTRTVACAGCARCKGPLETDCHEQCAAGCTGPRHSDCLAC 255
QY 275 LHFNSGICELHCPLATVYNTDFFESMPNPEGATGASCVTACPNYXSTDVGS 329
Db 256 LHFNSGICELHCPLATVYNTDFFESMPNPEGATGASCVTACPNYXSTDVGS 315
QY 330 ----- 329
Db 316 PLINQEVTAEDGTQRCCKPCARVCYGLGMEHLREVRANVTSANIQEFACCKIKGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLOPEQLQVETLEITGYLIISAMPDSLPLSVFQNLQVINGRI 435
QY 330 ----- 329
Db 436 LHNAGVSLTGLGISMGLRSRLRELGSLALIHNTHLCEVHTVPWDLFRNPQALH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQCLARGHCMPGPTQCVCNCSQFLRGCEVECHVLOGLPREY 555
QY 330 ----- 329
Db 556 NARHCLPCHPECCOPONGSVTCFGEPEADOCVACAHRKDPFCVARCPSGVKPDLSYMPWK 615
QY 330 ----- 329
Db 616 FPDEBACQCPINCTHSCVDLDDKCCPABQRASPLTISIISAVGILLVVLGVGEGILI 675
QY 330 ----- 329
Db 676 KRQOKIRKTYMRRLQETELVEPLTPSGAMPQAOQMRILKETELRKVKVLSGARFTY 735
QY 330 ----- 329
Db 736 KGIWPDGENVKIPVAIKVLENTSPRANKELIDEAVYVAGSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCGLDHRVBNRGLSODLLMKMOIAKMSYLEVRLVHRDLAARNVLYKS 855
QY 330 ----- 329
Db 856 PNHVKTIDFGLARLDIDETEHADGKVPKIMMALESILRRFTHQSDVMSGYVWEL 915
QY 330 ----- 329
Db 916 MTEGAPYDGIPIAREIPDLLEKGERLPPEICTIDVYIMVKCMTIDSECRPRFRELVE 975
QY 330 ----- 329

```



```

Db 976 FSRMARDPQRFVYIQLNEDLGPASPLDSTFYRSLEDDMDLVDAEELVPOQGFCCPD 1035
Qy 330 --GAGVWHHRHRSSTRSGGDLTLGLEPSEEPASPLAPSGAGSDVFDGLGAA 387
Db 1036 APGAGVWHHRHRSSTRSGGDLTLGLEPSEEPASPLAPSGAGSDVFDGLGAA 1095
Qy 388 KGIQSLPTHDSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEVNOPDVROPSPREG 447
Db 1096 KGIQSLPTHDSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEVNOPDVROPSPREG 1155
Qy 448 PLPARAGATLERAKTLSSGKNGVADVAFAGAVENPEYLTPQGAAPQHPPPAFSP 507
Db 1156 PLPARAGATLERAKTLSSGKNGVADVAFAGAVENPEYLTPQGAAPQHPPPAFSP 1215
Qy 508 AFDLYWDDPPERGAAPSTFKGTPTAENPEYGLDVP 546
Db 1216 AFDLYWDDPPERGAAPSTFKGTPTAENPEYGLDVP 1254

```

RESULT 12

```

US-09-821-883-5
; Sequence 5: Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 479
; TYPE: PR
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER300*-rGM-CSF construct
US-09-821-883-5

```

Query Match 58.5%; Score 2188.5; DB 10; Length 479;

Best Local Similarity 61.7%; Pred. No. 3.3e-121; Indels 211; Gaps 4;

Matches 426; Conservative 22; Mismatches 31;

```

Qy 1 MRAAPLLARASLSLGFLEFLFWLDRSVLAKELARGAASSTQVCTGTDKRLRPASPET 60
Db 1 MRAAPLLARASLSLGFLEFLFWLDRSVLAKELARGAASSTQVCTGTDKRLRPASPET 60
Qy 61 HLDMLRLYGGCOVVOGNLELTYLPTNASLSFLDIOEVOGYVLIANNQVROPVLRIRI 120
Db 61 HLDMLRLYGGCOVVOGNLELTYLPTNASLSFLDIOEVOGYVLIANNQVROPVLRIRI 120
Qy 121 VRGTQLEEDNVYALAVLNGDPPLNNTVTGASPGELRELDRLSTELLKGGVLIQRNPOL 180
Db 121 VRGTQLEEDNVYALAVLNGDPPLNNTVTGASPGELRELDRLSTELLKGGVLIQRNPOL 180
Qy 181 CYODTILMKDIFHKNNQALTLIDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRVCA 240
Db 181 CYODTILMKDIFHKNNQALTLIDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRVCA 240
Qy 241 GGCARCGRLPTDCCHQCAAGCTGPKHSDCLACLHFNHSGICELCPALVYNTDTFES 300
Db 241 GGCARCGRLPTDCCHQCAAGCTGPKHSDCLACLHFNHSGICELCPALVYNTDTFES 300
Qy 301 MPNEGRYTGASCVTACPNYLTSDVSGAGVWHHRHRSSTRSGGDLTLGLEPSEE 360
Db 301 MPNEGRYTGASCVTACPNYLTSDVSGAGVWHHRHRSSTRSGGDLTLGLEPSEE 360
Qy 361 EAPRSPLAPSEAGSDVFDGLGAAKGLQSLPTHDSPLQRYSEDPVPLPSETDGYV 420

```

```

Db 332 ----- 331
Qy 421 APLTCSQPEVNOPDVROPSPSPREGRLPARAGATLERAKTLSSGKNGVADVAFAG 480
Db 332 ----- 334
Qy 481 GAVENPEYLTPQGAAPQHPPPAFSPADNLYWDDPPERGAAPSTFKGTPTAENPEY 540
Db 335 ----- 338
Qy 541 LGLDVPAARASPSSTOPWENHVALIQEARLLNLSRDTAENKETEVEISEMFDLOEP 600
Db 339 ----- 391
Qy 601 TCIQTRLEIKKQGRSLTLYKGLPTKMAASHYKONCPPTBETSCATQIITFESKELKD 660
Db 392 TCQVTRLEIKKQGRSLTLYKGLPTKMAASHYKONCPPTBETDELEVTTFEDEFKMLKG 451
Qy 661 FLVYIPDCWEPVOEGAPPPAAHHNNHH 690
Db 452 FLVYIPDCWEPVOEGAPPPAAHHNNHH 479

```

RESULT 13

```

US-09-854-356-14
; Sequence 14: Application US/09854356
; Patent No. US20020177567A1
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Gheysen, Dirk
; APPLICANT: Corixa Corporation
; APPLICANT: SmithKline Beecham Biologicals S. A.
; TITLE OF INVENTION: HER-2/neu Fusion Proteins
; FILE REFERENCE: 014058-009810PC
; CURRENT APPLICATION NUMBER: US/09/854,356
; PRIOR FILING DATE: 2001-05-09
; PRIOR APPLICATION NUMBER: US 09/493,480
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: US 60/117,976
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1256
; TYPE: PR
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse HER-2/neu protein
US-09-854-356-14

```

Query Match 52.4%; Score 1961; DB 9; Length 1256;

Best Local Similarity 54.3%; Pred. No. 2.4e-107; Indels 740; Gaps 3;

Matches 434; Conservative 23; Mismatches 67;

```

Qy 11 AASLSLGFLEFLFWLDRSVLAKELARGAASSTQVCTGTDKRLRPASPTHMLRLYQ 70
Db 4 AAMCRMGFLAL-----LSPGAAGTQVCTGTDKRLRPASPTHMLRLYQ 51
Qy 71 GCOVVOGNLELTYLPTNASLSFLDIOEVOGYVLIANNQVROPVLRIRIYRGTQLEEDN 130
Db 52 GCOVVOGNLELTYLPTNASLSFLDIOEVOGYVLIANNQVROPVLRIRIYRGTQLEEDN 111
Qy 131 YALAVIDNGPPLNNTVTGASPGELRELDRLSTELLKGGVLIQRNPOLCTODITLTK 169
Db 112 YALAVIDNGPPLNNTVTGASPGELRELDRLSTELLKGGVLIQRNPOLCTODITLTK 171
Qy 190 DIFHKNNQALTLIDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRVYAGGARCKGP 249
Db 172 DVLKNNQALPVMMDTNRSRACHPCSPMKCSRCWGESSEDCOSLTRVYAGGARCKGP 231
Qy 250 LPTDCHQCAAGCTGPKHSDCLACLHFNHSGICELCPALVYNTDTFESMPNPEGRYT 309

```

Db	232	LPTDCHBOCAAGCGPRKHSCLACLFHNSGICELCPALITYNTDTFSMLNPGRYT	291
Qy	310	FGASVTCAPVNNYSTDGS-----	329
Db	232	FGASVTCAPVNNYSTDGS-----	329
Qy	330	-----	329
Db	352	RGARATISDNIQEFAGCKKIRGSLAFLPESPDPNPSSGAPLKRPHQVETLEETGYL	411
Qy	330	-----	329
Db	412	YISANPESEFQDLVYFQNLRIYGRILHDGAYSLTLQGLGHSGLRSLRELGSGLALHHR	471
Qy	330	-----	329
Db	472	NTHLCFVNTVPMDQLFRNPQALLHSGNRPEACGLEGLYCNSLCARGHCKMGCPPTQCVN	531

```

: Sequence 118, Application US/09870759
: Patent No. US20020177551A1
:
: GENERAL INFORMATION:
: APPLICANT: TERMAN, David S
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
: FILE REFERENCE: 870759
: CURRENT APPLICATION NUMBER: US/09/870,759
: CURRENT FILING DATE: 2002-01-14
: PRIOR APPLICATION NUMBER: US 60/208,128
: PRIOR FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 166
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 118
: LENGTH: 1260
: TYPE: PRY
: ORGANISM: Rattus norvegicus
:
: US-09-870-759-118

```

Query Match	52.2%	Score 1950.5	DB 9	Length 1260
Best Local Similarity	34.1%	Pred. No. 9.8e-107		
Matches 433	Conservative 24	Mismatches 72	Indels 741	Gaps 3

0Y	330	-----	329
0Y	330	-----	329
Db	592	KDSSCVARCPBGVCPDLSTYMPIMKYPDDEBICQPCPLNCHTSCVDLDERCGPAEQRAS	651
0Y	330	-----	329
Db	652	VFIATVGVLLFLIIVVIGILIKRRRQKIRKTYMRLLQETELVEPLTPSGAVPQA	711
0Y	330	-----	329
Db	712	QMRILETELKRLKVLGSGAGTGYKGIWIPDGENVKIPVAIKVLENTSPRANKETLDE	771
0Y	330	-----	329
Db	772	AVVAGVGSPPYVSRLLGICLTSTVOLYQMLPYGCLLDHVHREHGRGLSODLIMKVOIA	831
0Y	330	-----	329
Db	832	KCMXYLEEVRLVHRDLAARNVLVKSNNHVKITDFGLARLIDIDETEHADGKVP	891
0Y	330	-----	329
Db	892	LESILRRRTHOSDVWYSGVTVWELMTFGANPYDGIIPAREIPDLLEKGERLP	951
0Y	330	-----	329
Db	952	VYMIWVKCMIDSECRPRFRELYSEFSRMANDPQRFVYIQONEDLGSSPMOSTFYRSLLE	1011
0Y	330	-----	362
Db	1012	DDDMGELVDAEELVYPOOGFSPDPLCTGSTAHRRHRSSSARSGGELTGLLEPSEEP	1071
0Y	363	PRSPAPSEBGASDVEDDLGAKKGLQSLPTHDSPLQRYSEDPYPLPSENDGYAP	422
Db	1072	PRSPAPSEBGASDVEDDLGAVTKGLQSLPHDLSPLQRYSEDPYPLPPEPDGYAP	1131
0Y	423	LTCSPQPEYVNOVDVPRQPPSPBREGPLPAARPAAGATLERAKTILSGKNGVKDVAFAGCA	482
Db	1132	LACSPQPEYVNOVDVPRQPPSPBREGPLPAARPAAGATLERAKTILSGKNGVKDVAFAGCA	1191
0Y	483	VENPEYLTPOGGAPOHPHPPAFSPAFDNLVYWDOPPERGAPPSTFGKPTAENPEYL	542
Db	1192	VENPEYLTPOGGAPOHPHPPAFSPAFDNLVYWDONSSSEGGPPSTFGKPTAENPEYL	1251
0Y	543	LDVP	546
Db	1252	LDVP	1255
RESULT 14			
US-09-870-759-118			

```

QY 330 ----- 329
Db 770 KEIDEAVYMGVGSPIYVSRLLGICLTSTVOLVQIMPYGCLLDVREHGRGLSSQDLN 829
QY 330 ----- 329
Db 830 WCVQAKMSYLEIDVRLVHRDLAARNVLKSPNHVKITDFGLARLLDIDETEHADGSKV 889
QY 330 ----- 329
Db 890 PIKMALESLRRRTFQSDVMSYGVTVMEIMTEGAKPYDGIPIAREIPDLLEKGERLPQ 949
QY 330 ----- 329
Db 950 PICTIDVYIMVWKCMIDSECRPRRELVSFESRMARDPQRFVYIQNEDLGSSPMDSF 1009
QY 330 ----- 356
Db 1010 YRSLLEDMDGDLVDAEEYLVPOGFSPDPTPGTGSTAHRHRSSSTRSGGELTGLE 1069
QY 337 PSEEAAPSPAPSPGASDVFDGDLGAKGLQSLPTHPSPLOKRSDEPTVLPSET 416
Db 1070 PSEEBPPSPAPSPGASDVFDGDLGAVTKGLQSLSPHLSPLQKRSDEPTVLPSET 1129
QY 417 DGYPALPCSPQPEYVNPQVPPSPREGPLPAAPAGATLERAKTILSPKNGVQV 476
Db 1130 DGYPALPCSPQPEYVNPQVPPSPREGPLPAAPAGATLERAKTILSPKNGVQV 1189
QY 477 FAFGAVENPEYLVPOGGAAPPPAPSPAFDNLVYWDOPPERGAPSTFKGTPTAE 536
Db 1190 FAFGAVENPEYLVREGTASPPHSPAFDNLVYWDONSSBOGPPSPNFEETPTAE 1249
QY 537 NPEYGLDVP 546
Db 1250 NPEYGLDVP 1259

```

RESULT 15

US-09-854-356-2

Sequence 2, Application US/09854356

Patent No. US2002017367A1

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Gheysen, Dirk

APPLICANT: Corixa Corporation

APPLICANT: SmithKline Beecham Biologicals S. A.

TITLE OF INVENTION: HER-2/neu Fusion Proteins

FILE REFERENCE: 014058-009810PC

CURRENT APPLICATION NUMBER: US/09/854,356

PRIOR APPLICATION NUMBER: 2001-05-09

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: US 60/117,976

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 1256

TYPE: PRT

ORGANISM: Rattus sp.

FEATURE:

OTHER INFORMATION: rat HER-2/neu protein

NAME/KEY: DOMAIN

LOCATION: (1)..(654)

OTHER INFORMATION: extracellular domain (ECD)

NAME/KEY: DOMAIN

LOCATION: (677)..(1256)

OTHER INFORMATION: intracellular domain (ICD)

NAME/KEY: DOMAIN

LOCATION: (771)..(998)

OTHER INFORMATION: kinase domain (KD)

NAME/KEY: DOMAIN

```

LOCATION: (991)..(1256)
OTHER INFORMATION: phosphorylation domain (PD)
NAME/KEY: DOMAIN
LOCATION: (991)..(1049)
OTHER INFORMATION: fragment of the phosphorylation domain, preferred
OTHER INFORMATION: portion (delta PD)
US-09-854-356-2

```

```

Query Match 52.1% Score 1948: DB 9: Length 1256;
Best Local Similarity 34.3% Pred. No. 14e-166;
Matches 433: Conservative 21: Mismatches 70: Indels 740: Gaps 3:

```

```

QY 11 AASLIGFLFLFEMLDRLSLAKELARGASTOYCTGDMKRLPASPETHLMDLRLHQ 70
Db 4 AAMCRMGFLAL-----LPPGIAGTOYCTGDMKRLPASPETHLMDLRLHQ 51
QY 71 GCQVVOGNLELYLPTNASLSFLQDIQEVQGYLVIANQVQVPLQRLRYRGTOLEFDN 130
Db 52 GCQVVOGNLELYLVANASLSFLQDIQEVQGYLVIANQVQVPLQRLRYRGTOLEFDK 111
QY 131 YALAVLDGDDPLNNTPTV-GASPGLEQLQSLFTELKGVILQRRPOLCYQDTLWK 189
Db 112 YALAVLDGDDPLNNTPTV-GASPGLEQLQSLFTELKGVILQRRPOLCYQDTLWK 171
QY 190 DIFHKNQALATLIDTNRSRACHPCSPMKGSRGSESSSEDCOSLRTVACAGCARKGP 249
Db 172 DIFHKNQALATLIDTNRSRACHPCSPMKGSRGSESSSEDCOSLRTVACAGCARKGP 231
QY 250 LPTDCHEQCAAGCTGPRHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNREGRT 309
Db 232 LPTDCHEQCAAGCTGPRHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNREGRT 291
QY 310 FGASCVTCAPNYLSTDVGS----- 329
Db 292 FGASCVTCAPNYLSTDVGS----- 351
QY 330 ----- 329
Db 352 RGAARITSDNVOEPDGCCKIFGSLAFLPESFDGDPSSGIAPLRPQQLVFTLEETGYL 411
QY 330 ----- 329
Db 412 YISAMPDSLRLSVFQNLRIIRGRILHDGAYSLTLOGIGHSLSRLSRELGSGLALILR 471
QY 330 ----- 329
Db 472 NAHLCTVHTVMDOLFRRPHQALLHSGNRPEBDCLEGLVCSNLCAHGCWGPPTQCVN 531
QY 330 ----- 329
Db 532 CSHPLRGQCEVEECVWKGGLPREVYSKRCULPCHPEQOPONSSETCFESBADQCAACAHY 591
QY 330 ----- 329
Db 592 KDSSSCVAPRCPGKPDLSYMPIMKYPDEEGICOPCPICTHSCVDLDERGCPAPORASP 651
QY 330 ----- 329
Db 652 VTIIATVEGYLLFLILVYVVGILIKRRRQKIRKTYMRLLQETELVPLIPSGAMPNQA 711
QY 330 ----- 329
Db 712 QMRILKETELRKVKVLSGAGFYVKGIMIPDGENVKIPVAKVLRENTSFRANKELDE 771
QY 330 ----- 329
Db 772 AYVAVGSPYVSRLLGICLTSTVOLVQIMPYGCLLDVREHGRGLSSQDLNMCVQIA 831
QY 330 ----- 329
Db 832 KMSYLEIDVRLVHRDLAARNVLKSPNHVKITDFGLARLLDIDETEHADGSKVPIKMA 891
QY 330 ----- 329

```

```

Db 892 LESILRRRTHQSDVMSYGVTVWELMTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITID 951
QY 330 ----- 329
Db 952 VYMIMVCMKIDSECRPRFRELVSFBRMARDPQRFVVIQNEIDLGPSSPMDSTFYRSLE 1011
QY 330 -----GACGMVHHRRSSSTRSGGDLTLGLEPSEEA 362
Db 1012 DDDMGDLVDAEEXLVPOQGFSPDPPTGTGTAHRRHSSSTRSGGELTLGLEPSEEGP 1071
QY 363 PRSPLAPSEAGSDVFPDGLGMAKGLQSLPTHDPSPLOQRYSDPPTVPLPSETDGYVAP 422
Db 1072 PRSPLAPSEAGSDVFPDGLGMAKGLQSLPTHDPSPLOQRYSDPPTVPLPSETDGYVAP 1131
QY 423 LTCSPQPEYVNOPDVRPQPPSPREGPLPAARPAAGATLERAKTLPCKNGVVKDVFAGGA 482
Db 1132 LACSPQPEYVNOSEVQPPPLTPEGPLPYVRPAGATLERPKTLPCKNGVVKDVFAGGA 1191
QY 483 VENPEYLTPOGGAAPOPHPAPAFSPAFDNLXYMDODPPERGAPPSTFKGPTAENPEYLG 542
Db 1192 VENPEYLVPREGTASPPHPAPAFSPAFDNLXYMDQNSSEOGPPPSNEGTPTAENPEYLG 1251
QY 543 LDVP 546
Db 1252 LDVP 1255

```

Search completed: April 28, 2003, 13:44:03
 Job time : 40.3846 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:39:54 ; Search time 20.1923 Seconds
(without alignments)
1005.423 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLIARASTISLGLF.....EPVOEGAPPPAAHHHHH 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfill1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	1 US-08-467-083-68	Sequence 68, Appl
2	2405.5	64.3	1255	1 US-08-414-417B-68	Sequence 68, Appl
3	2405.5	64.3	1255	2 US-08-486-348A-68	Sequence 68, Appl
4	2405.5	64.3	1255	2 US-08-625-101-2	Sequence 2, Appl
5	2405.5	64.3	1255	2 US-08-468-545B-68	Sequence 68, Appl
6	2405.5	64.3	1255	2 US-08-356-786-2	Sequence 2, Appl
7	2405.5	64.3	1255	3 US-08-466-680B-68	Sequence 68, Appl
8	2385.5	63.8	1255	2 US-08-484-438-8	Sequence 8, Appl
9	2224	59.5	782	3 US-09-146-283-4	Sequence 4, Appl
10	2224	59.5	782	3 US-08-579-823A-4	Sequence 4, Appl
11	2224	59.5	782	4 US-09-344-195-4	Sequence 4, Appl
12	1607	43.0	419	4 US-09-630-155-2	Sequence 2, Appl
13	1597	42.7	624	3 US-08-422-108-1	Sequence 1, Appl
14	1597	42.7	624	4 US-08-422-734-1	Sequence 1, Appl
15	1183	31.6	580	1 US-08-414-417B-69	Sequence 69, Appl
16	1183	31.6	580	2 US-08-486-348A-69	Sequence 69, Appl
17	1183	31.6	580	2 US-08-468-545B-69	Sequence 69, Appl
18	1183	31.6	580	3 US-08-466-680B-69	Sequence 69, Appl
19	761	20.4	515	2 US-09-146-283-2	Sequence 2, Appl
20	761	20.4	515	3 US-08-579-823A-2	Sequence 2, Appl
21	761	20.4	515	4 US-09-344-195-2	Sequence 2, Appl
22	721	19.3	644	1 US-08-336-708A-9	Sequence 9, Appl
23	721	19.3	911	2 US-08-484-438-10	Sequence 10, Appl
24	721	19.3	1058	2 US-08-484-438-4	Sequence 7, Appl
25	721	19.3	1210	2 US-08-484-438-7	Sequence 4, Appl
26	721	19.3	1210	2 US-08-475-035-4	Sequence 2, Appl
27	721	19.3	1308	2 US-08-484-438-2	Sequence 2, Appl

28	704.5	18.8	478	4 US-09-570-454-2	Sequence 2, Appl
29	698.5	18.7	1343	6 5183884-4	Patent No. 5183884
30	697	18.6	1342	1 US-07-978-895-4	Sequence 4, Appl
31	697	18.6	1342	2 US-08-484-438-9	Sequence 9, Appl
32	697	18.6	1342	2 US-08-473-119-4	Sequence 4, Appl
33	697	18.6	1342	2 US-08-475-352-4	Sequence 4, Appl
34	693	18.5	274	3 US-08-469-318-144	Sequence 144, App
35	693	18.5	274	3 US-08-468-609A-144	Sequence 144, App
36	693	18.5	274	4 US-08-446-872A-144	Sequence 144, App
37	693	18.5	274	4 US-08-762-227A-144	Sequence 144, App
38	693	18.5	274	5 PCT-US93-01185-144	Sequence 144, App
39	680.5	18.2	301	3 US-08-469-318-142	Sequence 142, App
40	680.5	18.2	301	3 US-08-468-609A-142	Sequence 142, App
41	680.5	18.2	301	4 US-08-446-872A-142	Sequence 142, App
42	680.5	18.2	301	4 US-08-762-227A-142	Sequence 142, App
43	680.5	18.2	301	5 PCT-US93-01185-142	Sequence 142, App
44	679	18.2	165	1 US-08-318-193-8	Sequence 8, Appl
45	678	18.1	168	1 US-08-318-193-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-467-083-68
Sequence 68, Application US/08467083
Patent No. 5726023
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/NEU ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,083
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/414,417
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TELEX: 3723836 SEDANBERRY
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-467-083-68
Query Match 64.3%; Score 2405.5; DB 1; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;
35 LARGAASOVCTGTDMKRLPASPETHLDMLRHLYOGGVGVGNLELYLPTNASTLSFLQ 94

```

Db 16 LPPGAATGYCTGTDMLRLPASPEHLDMLRHLRYGCGVVOGNLELTYLPTNASLSFLQ 75
QY 95 DIEVOGYVLIANOVQVLOQLRLRYRGQLEFEDNALAVLNDGDLNNTTPTGASPG 154
Db 76 DIEVOGYVLIANOVQVLOQLRLRYRGQLEFEDNALAVLNDGDLNNTTPTGASPG 135
QY 155 GLRELQRLSLTEILKGVLQORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 214
Db 136 GLRELQRLSLTEILKGVLQORNPOLCYODTILMKDIFHKNNOLATLIDTNSRACHPC 195
QY 215 SPMCKGRGCGESSEDCOSLTRTYCAGGACARCKGRLPTDCHECAAGCTGPKHSDCLAC 274
Db 196 SPMCKGRGCGESSEDCOSLTRTYCAGGACARCKGRLPTDCHECAAGCTGPKHSDCLAC 255
QY 275 LHFNHSGICELHCPALVTYNTDFESMPNPEGRTTGASCVTACPYNYLSTDVGS 329
Db 256 LHFNHSGICELHCPALVTYNTDFESMPNPEGRTTGASCVTACPYNYLSTDVGSCTLYC 315
QY 330 329
Db 316 PLHNOEVTAEEDGQRCCKSPCARVCYGLGMEHLREAVTSANIQEFAGCKKIFGSLA 375
QY 330 329
Db 376 FLPESEFGDPAANTAPLQPOLOVFEETLEITGYLISAMPDLSPLDLSVFONQVIRGRI 435
QY 330 329
Db 436 LHNQAVSLTLQGLISWGLRSLRELGSGLALHNNTHLCFVHTVPMWDLFRNPHQALH 495
QY 330 329
Db 496 TANRPEDECVGEGLACHQLCARHCWGPPTQCVNCSQFLRGQCEVCECHVLOGLAPREV 555
QY 330 329
Db 556 MARHCLPHECOPQNSVTCFGEADQCAHAHYKDPFCVACRPSGVKPDLSYPMIMK 615
QY 330 329
Db 616 FPBEGACOPCINCHTSCVDLDDKGRPAQRASPLTISIISAVGILLVYVGLVFEILI 675
QY 330 329
Db 676 KRROQIRKXYTMRLLQETELVEPLTPSGAMPQAOHRIKETELRKVYKVGSGAFGY 735
QY 330 329
Db 736 KGIWIPDGENVKIPVAIKULRENTSPKANKELIDEAVYMAVGSPYVSRLLGICLSTVO 795
QY 330 329
Db 796 LVTQLMPYGLDLHVENRGRGLSQDILLNMCQIAKMSYLEDRVLRDLAARNVLVKS 855
QY 330 329
Db 856 PNHVKTIDFGLARLLDIDETEHADGKVPKIMMALESILRRRTHOSDVMSGYVMEL 915
QY 330 329
Db 916 MTRGAKPYDGIIPARELPDLLEKGBRLPPICITIDVYIMVCMIDSECRPRELVESE 975
QY 330 329
Db 976 FSHMARDPQRFVVIQNEDELGPASPLDSTFYRSLLEDDMDLDAEEXYLVPOGFECPDP 1035
QY 330 387
Db 1036 APGAGGVVHRRSSSTRSGGDLTLGLEPSEEDAPRSLAPSEGAGSDVFDGDLGAA 1095
QY 388 KGIQSLPTHPSPLOVRSDEPTVPLPSEMTGYAFLTCSQPEYVNPQDVRPQPPSPREG 447

```

```

Db 1096 KGIQSLPTHPSPLOVRSDEPTVPLPSEMTGYAFLTCSQPEYVNPQDVRPQPPSPREG 1155
QY 448 PLPARAGATLERAKTLSPGKNGVVDVAFGAVENPEYLTPOGGAQDPHPPPAFSP 507
Db 1156 PLPARAGATLERAKTLSPGKNGVVDVAFGAVENPEYLTPOGGAQDPHPPPAFSP 1215
QY 508 AFDNLYWDDPPPERGAPSPSTFKGTPTAENPEYLGIDVP 546
Db 1216 AFDNLYWDDPPPERGAPSPSTFKGTPTAENPEYLGIDVP 1254

RESULT 2
US-08-414-417B-68
; Sequence 68, Application US/08414417B
; Patent No. 5801005
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C2
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-414-417B-68

Query Match 64.3%; Score 2405.5; DB 1; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

```

```

|||||
Db 256 LHFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACPYNLTSDVSCSLVC 315
Qy 330 ----- 329
Db 316 PLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIOBFACKKIFGSLA 375
Qy 330 ----- 329
Db 376 FLPSFDGDPASNTAPLOPBOLOVFETLEETGYLYISAMPDSLPLDSVFONLOYIRGRI 435
Qy 330 ----- 329
Db 436 LHNCAVSLTLOGLGISWLGRLSRELGSGLALIHNNHLCFVHTVPMQDLFRNPQALLH 495
Qy 330 ----- 329
Db 496 TANRPEDECVGEGLACHOLCARCWCWBP7QCVNCSOFLRGBCVEECRVLOGLPREYV 555
Qy 330 ----- 329
Db 556 NARHCLPCHRECPONGSVTCFGEADOCVACAHYKDPFECVACPSGVKFDLSMPYWK 615
Qy 330 ----- 329
Db 616 FPDEGACQPCPINCTHSCVDLDDKCPAQRASPLTISIAYVIGILLVVLGVFGIL 675
Qy 330 ----- 329
Db 676 KRROOKIRKYMRLLOETELVEPLTPSGAMPNOQMRLKETELRKVKVLSGSAFGTVY 735
Qy 330 ----- 329
Db 736 KGIMIPOGENKITYAIVKILRENTSPRANKRILDEAYVMAGVSPYSRLGICLTSTVO 795
Qy 330 ----- 329
Db 796 LVTLMPYCLLDHYRENRGRIGSODLLMCMQIAKMSYLEVRLVHRDLAARNVLKVS 855
Qy 330 ----- 329
Db 856 PNHVKITDFGLARLLIDETERYHADGKVPKIMMALESILRRRPTHOSDWSYGVYMWEL 915
Qy 330 ----- 329
Db 916 MTFGAKPYDGTIPAREIPDLLEKGERLPQPICTIDVYIMVIMKCMIDSECRPRRELAYSE 975
Qy 330 ----- 329
Db 976 FSRMARDFQRENVIONEDLGASPLDSTFYRSLLLEDDMDGLVDAAEYLVPOGFCFDP 1035
Qy 330 --GAGGWHHRRSSSTFGGGDLTLGLEPSEDEAPRSLAPRSGAGSDVDGDLGMAA 387
Db 1036 APGAGGWHHRRSSSTFGGGDLTLGLEPSEDEAPRSLAPRSGAGSDVDGDLGMAA 1095
Qy 388 KGLSLPTHDSPLOKRYSEDPVPLPSETDGYVAPLTCSPQREYVNOVDVPRPPSPREG 447
Db 1096 KGLSLPTHDSPLOKRYSEDPVPLPSETDGYVAPLTCSPQREYVNOVDVPRPPSPREG 1155
Qy 448 PLPARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQPHRPAESP 507
Db 1156 PLPARPAGATLERAKTSLSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQPHRPAESP 1215
Qy 508 AFDNLVYWDODRPERGAPSPSTFKGPTAENPEYLGIDVP 546
Db 1216 AFDNLVYWDODRPERGAPSPSTFKGPTAENPEYLGIDVP 1254

```

RESULT 3
 US-08-486-348A-68
 Sequence 68 Application US/08486348A
 Patent No. 5846538
 GENERAL INFORMATION:

```

APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,348A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C6
TELEPHONE: (206) 682-4800
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-486-348A-68

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

Qy 35 LARGAASVOVCTGDMKRLDPAEPETHLDMRLHYOGCOVVOGNEUETLYLPTNASTSLQ 94
Db 16 LPPGAASTOVCTGDMKRLDPAEPETHLDMRLHYOGCOVVOGNEUETLYLPTNASTSLQ 75
Qy 95 DIOEVQGVYLLAHNOVROVPLORLRYVGTOLFEDNVYALAVLDNDPPLNNTPTPTGASPG 154
Db 76 DIOEVQGVYLLAHNOVROVPLORLRYVGTOLFEDNVYALAVLDNDPPLNNTPTPTGASPG 135
Qy 155 GLRELQLRSLTEILKGVLIQIORNPOLCYQDPTILMKDIFHKNNQALATLIDTNRSRACHPC 214
Db 136 GLRELQLRSLTEILKGVLIQIORNPOLCYQDPTILMKDIFHKNNQALATLIDTNRSRACHPC 195
Qy 215 SPMCKGSRKCESSESDCOSLTRTYCAGCAGCARKGRLPTDCCHEGCAAGCTGPKHSIDLAC 274
Db 196 SPMCKGSRKCESSESDCOSLTRTYCAGCAGCARKGRLPTDCCHEGCAAGCTGPKHSIDLAC 295
Qy 275 LHFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACPYNLTSDVSCSLVC 329
Db 256 LHFHNSGICELHCPALVYNTDFESMPNPGRYTFGASCVTACPYNLTSDVSCSLVC 315
Qy 330 ----- 329
Db 316 PLHNOEYTAEDGTORCEKSKPCARVCYGLGMEHLREVAVTSANIOBFACKKIFGSLA 375
Qy 330 ----- 329
Db 376 FLPSFDGDPASNTAPLOPBOLOVFETLEETGYLYISAMPDSLPLDSVFONLOYIRGRI 435
Qy 330 ----- 329
Db 436 LHNCAVSLTLOGLGISWLGRLSRELGSGLALIHNNHLCFVHTVPMQDLFRNPQALLH 495
Qy 330 ----- 329

```

```

Db 496 TANRDECEVGEGLACHQLCARCHCWGPGPTQCVCNSQFLRGQCEVECEVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPHECOPQNGSVTCFGEADQCAAHYKDPPECVARGPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEGACOPCPINCHSCVLDLDDKGPAPQASPLTISAVVGLLVVGLVGVGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKTYMRRLLOETELVEPLTPSGAMPNOAQRILKETELRKVKVJGSGAGCTVY 735
QY 330 ----- 329
Db 736 KGIWPDGENVKIPVAIKVIRENTSPKANKEILDEAYVMAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTOLMPYGCLLDHRENRGRGLSGODLLNMCQIAKMSYLEVRLVHRDLAARNVLKVS 855
QY 330 ----- 329
Db 856 PNHVKTDFGLARLLDIDETEHADGSKVPDKMALESILRRRPTHOSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTEGAKYDDIPAREITDLEKEGRLPQPPICITIDVYMIWKCMIDSECRPRELVE 975
QY 330 ----- 329
Db 976 FSRMARPORFVVYQNEIDLSPASPLDSTFYRSLLEDDEDMGLVDAEEYLVPOQGFCCPD 1035
QY 330 ----- 387
Db 1036 APGGAGVHRRHSSSTRSGGDLTGLPSEEARPSPLAPSGASDVFDGGLGAA 1095
QY 388 KGLQSLTHTDPSLQRYSEDPTVPLPSETDGYVALTCSPPQEVNODVVRPQPPSPREG 447
Db 1096 KGLQSLTHTDPSLQRYSEDPTVPLPSETDGYVALTCSPPQEVNODVVRPQPPSPREG 1155
QY 448 PLPARARAGATLERAKTLSEKNGVAVDAFGGAVENPEYLPQGGAAQPPHPPAFSP 507
Db 1156 PLPARARAGATLERAKTLSEKNGVAVDAFGGAVENPEYLPQGGAAQPPHPPAFSP 1215
QY 508 AFDNLVYWDODPPERGAPSTFKGTPAENPEYLGLDVP 546
Db 1216 AFDNLVYWDODPPERGAPSTFKGTPAENPEYLGLDVP 1254

```

```

RESULT 4
US-08-625-101-2
Sequence 2, Application US/08625101
Patent No. 5869445
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disls, Mary L.
TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/625,101
FILING DATE: 01-Apr-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C7
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-625-101-2

```

```

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGASTQVCTGDMKRLPASPETHLDMLRHLVQCGVVOGNTLELYLPTNALSFLQ 94
Db 16 LPGAASTQVCTGDMKRLPASPETHLDMLRHLVQCGVVOGNTLELYLPTNALSFLQ 75
QY 95 DIOEVQGVYLIANQVQVPLQRLRYRGVQLFEDNVALVNDGDPNLTPTTGASPG 154
Db 76 DIOEVQGVYLIANQVQVPLQRLRYRGVQLFEDNVALVNDGDPNLTPTTGASPG 135
QY 155 GLRELOLRSLTEILKGGVLIQORNPOLCYODTILMKDIFHKNNOLALTLITNSRACHPC 214
Db 136 GLRELOLRSLTEILKGGVLIQORNPOLCYODTILMKDIFHKNNOLALTLITNSRACHPC 195
QY 215 SPKMGSRGCESESEDCSLTRTVYAGAGCARCKGPLPTDCHEGCAAGCTGPKHSDCLAC 274
Db 196 SPKMGSRGCESESEDCSLTRTVYAGAGCARCKGPLPTDCHEGCAAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDFFESNPNEGRTTFCASCVTACPYVLTSDGVS 329
Db 256 LHFHNSGICELHCPALVTYNTDFFESNPNEGRTTFCASCVTACPYVLTSDGVSCTLYC 315
QY 330 ----- 329
Db 316 PLHNOEVTAEEDGTORCEKSKPCARVCYGLGMEHLREVRVTSANIOEFAGCKKIEGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGASNTAPLOPELOQVETLEETGYLYISAMPDSLPLDSVFQNLQYINGRI 435
QY 330 ----- 329
Db 436 LHNQAVSLTQGLGISWGLRSLRELGSGLALIHNTHTLCFVHTVPWDQLFRNPQALHL 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQLCARCHCWGPGPTQCVCNSQFLRGQCEVECEVLOGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPHECOPQNGSVTCFGEADQCAAHYKDPPECVARGPSGVKPDLSYMPIMK 615
QY 330 ----- 329
Db 616 FPDEGACOPCPINCHSCVLDLDDKGPAPQASPLTISAVVGLLVVGLVGVGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKTYMRRLLOETELVEPLTPSGAMPNOAQRILKETELRKVKVJGSGAGCTVY 735

```



```

QY 330 ----- 329
Db 736 KGIWIPDGENKIPVAIKVIRENTSPKANKRIIDEAIVMAGVSPYVSRLLGICITSTVQ 795
QY 330 ----- 329
Db 796 LVTLQMPYGLLDHVRENRLGSQDLLNMCMQIAKMSYLEVRLVHRLAARNLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEFYHADGKVPKIMMALESILRRRTHOSDVSYGVYMWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVCKMIDSECRPRELVS 975
QY 330 ----- 329
Db 976 FSRMARDPQREVYIIONEDLGPASPLDSTFYRSLLEDDMDGLVDAEYLVPOGFFCPDP 1035
QY 330 ----- 387
Db 1036 APGAGVHHRRSSSTRSGGDLTLGLEPSEEAAPRSLAPSGAGSDVVDGDLGMAA 1095
QY 388 KGLQSLPTHDPSPQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNOVDVPPPPREG 447
Db 1096 KGLQSLPTHDPSPQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNOVDVPPPPREG 1155
QY 448 PLPARAPGATLERAKTLSPGKNGVVDVFAFGAVENPEYLTPQGAAPRPPPARSP 507
Db 1156 PLPARAPGATLERAKTLSPGKNGVVDVFAFGAVENPEYLTPQGAAPRPPPARSP 1215
QY 508 AFDNLVYWDODPPRGAAPSPFCKTPTAENPEYLGIDVP 546
Db 1216 AFDNLVYWDODPPRGAAPSPFCKTPTAENPEYLGIDVP 1254

RESULT 5
US-08-468-545B-68
Sequence 68 Application US/08468545B
Patent No. 5876112
GENERAL INFORMATION:
APPLICANT: Cheever, Marlin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,545B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sharkkey, Richard G.
REGISTRATION NUMBER: 32,629
REFERENCE/DOCKET NUMBER: 920010.448C5
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 68:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 1255 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-468-545B-68

Query Match      64.3%; Score 2405.5; DB 2; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGAATQVCTGTDMKRLPASPETHLMDLRHLVYOGGVQVGNLELYLPTNASLSFLQ 94
Db 16 LPFGAATQVCTGTDMKRLPASPETHLMDLRHLVYOGGVQVGNLELYLPTNASLSFLQ 75
QY 95 DIOEVQGVYLLAHNOVQVPLQRLRYVGTQVLFEDNVALAVLDGDDPLNNTPTVTSAPG 154
Db 76 DIOEVQGVYLLAHNOVQVPLQRLRYVGTQVLFEDNVALAVLDGDDPLNNTPTVTSAPG 135
QY 155 GLREIQLRSLTEILKGVLLQIQRNPOLCYQDTILMKDIFHKNNQALATLITNRSRACHPC 214
Db 136 GLREIQLRSLTEILKGVLLQIQRNPOLCYQDTILMKDIFHKNNQALATLITNRSRACHPC 195
QY 215 SPMCKGSRKMGESSEDCSLTRTYVAGGACARCKGPLPTDCHEOCAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRKMGESSEDCSLTRTYVAGGACARCKGPLPTDCHEOCAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELCHALVYNTDFESMPNPRGRTFGASCTACPYNTLSTDVGS----- 329
Db 256 LHFHNSGICELCHALVYNTDFESMPNPRGRTFGASCTACPYNTLSTDVGSCTLYVC 315
QY 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSKPCARVCGGLGMEHLREVRATVANSIOEFAGCKKIFGSLA 375
QY 330 ----- 329
Db 376 FLPSFDDPASNAPLQPEQLQVFELEITYGLYISAMPDLSPLDSVFNQVIRGRI 435
QY 330 ----- 329
Db 436 LHNQAYSLTLOGLSISGLSLRELISGLALLIHNTHLCFVHYVPMQDLFRNHQALLH 495
QY 330 ----- 329
Db 496 TANRDECEVGEGLACHQLCARCHQMGPPQVNCQSLRGQCEVRCVIGLPPREV 555
QY 330 ----- 329
Db 556 NARHCLPCHPCQPNQNGSVTCGPEADQVACAHYKDPFCVAPRPSGVKPPDLSYPIWK 615
QY 330 ----- 329
Db 616 FPDEBGAQPCPINCTHSCVDLDKCPAEOBASPLTISIAVVGILLVYLVGVVFGILI 675
QY 330 ----- 329
Db 676 KRROKIRKTYMRRLQETELVEPLTPSGAMPNQAQMLKETELRKVYVLSGAFGVY 735
QY 330 ----- 329
Db 736 KGIWIPDGENKIPVAIKVIRENTSPKANKRIIDEAIVMAGVSPYVSRLLGICITSTVQ 795
QY 330 ----- 329
Db 796 LVTLQMPYGLLDHVRENRLGSQDLLNMCMQIAKMSYLEVRLVHRLAARNLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEFYHADGKVPKIMMALESILRRRTHOSDVSYGVYMWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVCKMIDSECRPRELVS 975

```

QY 330 ----- 339
 Db 976 FSHARDPQRFVVIQNEEDLGPASPLDSTFYRSLLEDDMDGLVDAEYLVPOQGFCDP 1035
 QY 330 --GAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMA 387
 Db 1036 APAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMA 1095
 QY 388 KGQSLPTHPSPLOKRSSEPTVPLPSEIDGYAPLTCSPQPEYVNOQDVRPQPSREG 447
 Db 1096 KGQSLPTHPSPLOKRSSEPTVPLPSEIDGYAPLTCSPQPEYVNOQDVRPQPSREG 1155
 QY 448 PLPAAPBAGATLERAKTLSPGKNGVYKDYAFGAVENPEYLPPOGGAAPQPPAPFSP 507
 Db 1156 PLPAAPBAGATLERAKTLSPGKNGVYKDYAFGAVENPEYLPPOGGAAPQPPAPFSP 1215
 QY 508 AFNLYVWODPPRCAPSTFKGTPTAENPEYLGIDVP 546
 Db 1216 AFNLYVWODPPRCAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 6

US-08-356-786-2

Sequence 2, Application US/08356786
 Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.
 APPLICANT: Oppermann, Hermann
 APPLICANT: Houston, L. L.
 APPLICANT: Ring, David B.
 TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
 STREET: Exchange Place, 53 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/356,786

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-356-786-2

Query Match 64.3%; Score 2405.5; DB 2; Length 1255;
 Best Local Similarity 41.1%; Pred. No. 4.8e-179;
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LANGAASSTGVCTGDMKRLRPASPEHLDMLRLHYOCQVVOGNLELTYLPTNASLSFLQ 94

Db 16 LPGAASSTGVCTGDMKRLRPASPEHLDMLRLHYOCQVVOGNLELTYLPTNASLSFLQ 75
 QY 95 DIOEVGVYLIHNOYRQVLOLRIRYRGTQLEFEDNALVLVNGDPINNTVTGASPG 154
 Db 76 DIOEVGVYLIHNOYRQVLOLRIRYRGTQLEFEDNALVLVNGDPINNTVTGASPG 155
 QY 155 GLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOALTLIDTNSRACHPC 214
 Db 136 GLRELOLRSLTEILKGGVLIQRNPOLCYODTILMKDIFHKNNOALTLIDTNSRACHPC 195
 QY 215 SPKCKSRCKWESSESDCQSLTRFYCAGGACRCGRLPTDCHCQCAAGCTGPRHSDCLAC 274
 Db 196 SPKCKSRCKWESSESDCQSLTRFYCAGGACRCGRLPTDCHCQCAAGCTGPRHSDCLAC 255
 QY 275 LHFNHSGICEHCPALVTYTTDTFEESMPNPEGRTFGASCVTACPNYLSYTDVGS 329
 Db 256 LHFNHSGICEHCPALVTYTTDTFEESMPNPEGRTFGASCVTACPNYLSYTDVGS 315
 QY 330 ----- 339
 Db 316 PLHNOEYTAEDGTORCEKSKPCARVCYGLGMELREAVTSAHQEFAGCKKIFGSLA 375
 QY 330 ----- 339
 Db 376 FLPESEGDGPASNTAPLOPQLOVFTLEITGYLYISAMPDLPDLVFNQVIRGRI 435
 QY 330 ----- 339
 Db 436 LHNQAVSLTQGLISWLGRLSRLSGSLALIHNTHLFCVHTVPMDFLRPHQALH 495
 QY 330 ----- 339
 Db 496 TANRPEDECVGEGLACHQLCARGHCWGPPTQCVNSQFLRGCEVECHVLQGLPREV 555
 QY 330 ----- 339
 Db 556 NARHCLPCHPECPONGSVTCFGEADQVCAHYADPPFCVRCPSGVKPDLSYPIWK 615
 QY 330 ----- 339
 Db 616 FPDEBGAQCPINCITHSCVDLDDKGPAPQASPLTISAVGILLVYLVGFI 675
 QY 330 ----- 339
 Db 676 KRRQKIRKYMRLLOETELVEPLTPSGAMPQAMRLKETELRKVVGSGAFGVY 735
 QY 330 ----- 339
 Db 736 KGIMIPDGENVKIPAIKVIRENTSPRANKELIDEAVYMAVGSPYVSLGICLSTVQ 795
 QY 330 ----- 339
 Db 796 LVTQLMPYGCGLDHVHNRGRGLSQDLLNMCQIANGMSYLEDVRLVHRDLAARVLYKS 855
 QY 330 ----- 339
 Db 856 PNHVKITDGLARLIDIDEYHADGKVPKIMMALESILRRFTHQSDVMSYGVYVWEL 915
 QY 330 ----- 339
 Db 916 MTEGAKPYDGIAPAREIPDLLEKGERLPPICTIDYIMVYKWMIMDSECRPFRELVS 975
 QY 330 ----- 339
 Db 976 FSHARDPQRFVVIQNEEDLGPASPLDSTFYRSLLEDDMDGLVDAEYLVPOQGFCDP 1035
 QY 330 --GAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMA 387
 Db 1036 APAGGVVHHRHSSSTRSGGDLTGLERSEEARSPAPSEAGSDVFDGDLGMA 1095
 QY 388 KGQSLPTHPSPLOKRSSEPTVPLPSEIDGYAPLTCSPQPEYVNOQDVRPQPSREG 447

Db 1096 KGLSLPTHDSPLORYSEDTVPPLSETDGVAPLTCSPQPEVYNQDVPARQPPSPREG 1155
Qy 448 PLPAARPAATLERAKTLSPKNGVVKDYFAFGAVENPEYLTPOGGAAPQHPHPAPSP 507
Db 1156 PLPAARPAATLERAKTLSPKNGVVKDYFAFGAVENPEYLTPOGGAAPQHPHPAPSP 1215
Qy 508 AFDNLTYWDODPPERGAAPPSTFKGPTAENPEYLGLDVP 546
Db 1216 AFDNLTYWDODPPERGAAPPSTFKGPTAENPEYLGLDVP 1254

RESULT 7
US-08-466-6808-68
; Sequence 68, Application US/084666808
; Patent No. 6075122
GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,6808
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C4
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 682-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-466-6808-68

Query Match 64.3%; Score 2405.5; DB 3; Length 1255;
Best Local Similarity 41.1%; Pred. No. 4.8e-179;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

Qy 35 LARGAASOVCTGDMKRLPASPETHDMLRHLKOGCOVVOGNLELYLPTNASTLQ 94
Db 16 LPPGAASOVCTGDMKRLPASPETHDMLRHLKOGCOVVOGNLELYLPTNASTLQ 75
Qy 95 DIOEVQGVYLLAHNOVROVPLQRLRIVRGTOLEEDNVALAVLNDNDPLNNTTPVTGASPG 154
Db 76 DIOEVQGVYLLAHNOVROVPLQRLRIVRGTOLEEDNVALAVLNDNDPLNNTTPVTGASPG 135
Qy 155 GLREIQLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNQLALTILIDNRSACHPC 214
Db 136 GLREIQLSLTEILKGVLIQRNPOLCYODTILMKDIFHKNNQLALTILIDNRSACHPC 195
Qy 215 SPMCKGSRGKWSSESDCSLTFTVCAGCARCKGLPTDCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRGKWSSESDCSLTFTVCAGCARCKGLPTDCHEQCAAGCTGPKHSDCLAC 255
Qy 275 LHFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGS----- 329

Db 256 LHFHNSGICELHCPALVTYNTDFESMPNPEGRTFGASCVTACPYNYLSTDVGSCTLYVC 315
Qy 330 ----- 329
Db 316 PLHNOEYTAEDGTQRCCKSPCARVCYGLMEHLREYRAVTSANIOEFACCKIIFGSLA 375
Qy 330 ----- 329
Db 376 FLPESEFGDASPNTAPLOPEQLOVFETLEITGTYLISAMPDLSLPDLSTVONQYIRGRI 435
Qy 330 ----- 329
Db 436 LHNQAVSLTLQGLISWGLSLRELKSGSLAIHHNTHLCFVHTVPMQDLFRNPHQALLH 495
Qy 330 ----- 329
Db 496 TANREDECEVGEGLACHOLARCHCWGPPTQVCNCSQFLRGQCEVCRVLOGLPREYV 555
Qy 330 ----- 329
Db 556 NARHCLPCHPECOFONGSVTGFGEADQVACAHYKDPFCVAPCRPSGVKRPDLSTYMTWK 615
Qy 330 ----- 329
Db 616 FPDEGACQPCPINCTHSCVDLDDKGPAPQASPLTSIIASAVGILLVVLGVVFGILI 675
Qy 330 ----- 329
Db 676 KRQOKIRKYTMRLLOETELVEPLTPSGAMPNOAMRILKETELKVKVLSGSAFGTVY 735
Qy 330 ----- 329
Db 736 KGIWIPDQENKIPYAIKVLRENTSPKANKELDEAYVMAGVGSRYVRLIGICTSTVQ 795
Qy 330 ----- 329
Db 796 LVTOLMPYCLLDHYRENHGRIGSODLLNMCQIAKMSYLEDEVRLVHRDLARNVLVKS 855
Qy 330 ----- 329
Db 856 PNHYKITDFGLARLLDIDETEVHADGKVPKIMMALESILRRFTHOSDWSYGVTVWL 915
Qy 330 ----- 329
Db 916 MTFGAKPYDGIAPAREIPDLLEKGERLPPICTIDVYIMYKCMNIDSECRPRRELAVSE 975
Qy 330 ----- 329
Db 976 FSRMARDPQRFVLIQNEIDLGPASPLDSTFYRSLLLEDDMGDLVDAEYLVPPQGFPCDP 1035
Qy 330 --GAGGVNHHRRSSSTSGGGDLTLGLEPSEEAAPRSLPAPSBAGSDVDDGDMKAA 387
Db 1036 APGAGGVNHHRRSSSTSGGGDLTLGLEPSEEAAPRSLPAPSBAGSDVDDGDMKAA 1095
Qy 388 KGLSLPTHDSPLORYSEDTVPPLSETDGVAPLTCSPQPEVYNQDVPARQPPSPREG 447
Db 1096 KGLSLPTHDSPLORYSEDTVPPLSETDGVAPLTCSPQPEVYNQDVPARQPPSPREG 1155
Qy 448 PLPAARPAATLERAKTLSPKNGVVKDYFAFGAVENPEYLTPOGGAAPQHPHPAPSP 507
Db 1156 PLPAARPAATLERAKTLSPKNGVVKDYFAFGAVENPEYLTPOGGAAPQHPHPAPSP 1215
Qy 508 AFDNLTYWDODPPERGAAPPSTFKGPTAENPEYLGLDVP 546
Db 1216 AFDNLTYWDODPPERGAAPPSTFKGPTAENPEYLGLDVP 1254

RESULT 8
US-08-484-438-8
; Sequence 8, Application US/08484438
; Patent No. 581098
; Patent No. 581098 5780031

```

: GENERAL INFORMATION:
: APPLICANT: Plowman, Gregory D.
: APPLICANT: Cuioussou, Jean-Michel
: APPLICANT: Shoyab, Mohammed
: APPLICANT: Slegall, Clay B.
: APPLICANT: Hellster m, Ingegerd
: APPLICANT: Hellster m, Karl E.
: TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennle 6 Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,438
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/323,442
: FILING DATE: 14-OCT-1994
: APPLICATION NUMBER: US 08/150,704
: FILING DATE: 10-NOV-1993
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/981,165
: FILING DATE: 24-NOV-1992
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 5624-230
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ. ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-484-438-8

```

Query Match 63.8%; Score 2385.5; DB 2; Length 1255;
 Best Local Similarity 41.0%; Pred. No. 1.8e-177;
 Matches 509; Conservative 0; Mismatches 2; Indels 729; Gaps 3;

```

QY 35 LARGASTQVCTGDMKRLPASPETHLDMRLHLYGCGVYVGNLELYLPTNASLSFLQ 94
DB 16 LPGAASSTQVCTGDMKRLPASPETHLDMRLHLYGCGVYVGNLELYLPTNASLSFLQ 75
QY 95 DIDEVQGYVLIANNOVQVLOLRRLRYRGQTLEFEDNVALVNDGDPDLNNTTPTGASPG 134
DB 76 DIDEVQGYVLIANNOVQVLOLRRLRYRGQTLEFEDNVALVNDGDPDLNNTTPTGASPG 135
QY 155 GLRELQRLSTLEILKGVLLQORNPOLCYOPTILMKDIFHKNNOALATLIDTNSRACHPC 214
DB 136 GLRELQRLSTLEILKGVLLQORNPOLCYOPTILMKDIFHKNNOALATLIDTNSRACHPC 195
QY 215 SPMCKSGRCMGESSEDCQSLRTFVCAAGCARCGPLPTDCBHQCAAGCTGPKHSDCLAC 274
DB 196 SPMCKSGRCMGESSEDCQSLRTFVCAAGCARCGPLPTDCBHQCAAGCTGPKHSDCLAC 255
QY 275 LHRNHSICELHCPALVTYITDTFESMPNDEGRITFGASCVTACPYNYLSTDVGS----- 329

```

```

DB 256 LHRNHSICELHCPALVTYITDTFESMPNDEGRITFGASCVTACPYNYLSTDVGSCTLYC 315
QY 330 ----- 329
DB 316 PLHNOEVTAEADGTORCEKCKPCARVCYGLGMEHLREVRVAVTSANIOEFACCKIRGSLA 375
QY 330 ----- 329
DB 376 FLPEFSDGDPASMTAPLQRPQLOVFETLEITGYLYISAMPDSLPLDSVFNQVIRGRI 435
QY 330 ----- 329
DB 436 LHNQAVSLTLOGIGISWLGRLSRELGSGLALIHNTHLCFVHTVPMDQLFRNHQALLH 495
QY 330 ----- 329
DB 496 TANRPEDEVCEGLACHQLCARRALLGSGPTQCVNCSQFLRGQCYECCAVDGLPREYV 555
QY 330 ----- 329
DB 556 NARHCLPCHPECPQNGSVYCFGEADQVCAHAHYKDPFCVAPCPGVAKPDLSYMPIWK 615
QY 330 ----- 329
DB 616 FPDEGACOPCPINCHSCVLDLDCGCPABQASPLTISAVVGLLVVGLGVGIL 675
QY 330 ----- 329
DB 676 KRROQKIRKTMRLDQETELVEPLTPSGAMPNDAQMRILKETELRKVKVLSGAGCTYV 735
QY 330 ----- 329
DB 736 KGIWIPDGENVKIPVAKIVRENTSPKANKEILDEAYVAGVSPYVSLGICLSTVQ 795
QY 330 ----- 329
DB 796 LVTQLMPYGLLDHVRNKRGLSQDLLNKCQIANKMSTLEDVRLVHRLAARNLVKS 855
QY 330 ----- 329
DB 856 PNHVKITDFGLRLDIDETEVHADGKVPKMMALESILRRRTHQSDVMSVGYVMEL 915
QY 330 ----- 329
DB 916 MTEGAKPYDGPAREIPDLLEKGERLPQPICTIDVYIMVKCMMIDSECRPRRELVE 975
QY 330 ----- 329
DB 976 FSRMARDPQRFVYIQNEDLPASPLDSTFYRSLLEDMDKDLVDAEYLYVPOGFFCPDP 1035
QY 330 -GAGGAVHHRHSSSTRSGGDLITGLERPEEBAPRSLAPSEGAGSDVFDGLGAA 387
DB 1036 APGAGGVHHRHSSSTRSGGDLITGLERPEEBAPRSLAPSEGAGSDVFDGLGAA 1095
QY 388 KGIQSLPTHPSPLOKRSSEPTVPLPSETDGYVAPRLTCSQPEYVNOVDVRFQPPSPREG 447
DB 1096 KGIQSLPTHPSPLOKRSSEPTVPLPSETDGYVAPRLTCSQPEYVNOVDVRFQPPSPREG 1155
QY 448 PLPAARAGATLBRAKTLISGKNGVYKDVAFGAVENPEYLPPOGGAAPQPPRPFSP 507
DB 1156 PLPAARAGATLBRAKTLISGKNGVYKDVAFGAVENPEYLPPOGGAAPQPPRPFSP 1215
QY 508 AFDNLYWDDPPEBGAAPSTFKGTPT-AENPEYLGIDVP 546
DB 1216 AFDNLYWDDPPEBGAAPSTFKGTPVAVNPEY-GLDVP 1254

```

RESULT 9
 US-09-146-283-4
 : Sequence 4, Application us/09146283
 : Patent No. 5976546
 : GENERAL INFORMATION:

Query Match 59.5%; Score 2224; DB 3; Length 782;
 Best Local Similarity 59.2%; Pred. No. 3.6e-165;
 Matches 478; Conservative 18; Mismatches 104; Indels 208; Gaps 20;

```

OY 35 LARGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVQVGLLETTYPTNASTLSFLQ 94
DB 16 LPRGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVQVGLLETTYPTNASTLSFLQ 75
OY 95 DIOEVQGVYLIANNOVROVFLQRLRIYRGTOLEFEDNALAVLNDGDPINNTPTTGASPG 154
DB 76 DIOEVQGVYLIANNOVROVFLQRLRIYRGTOLEFEDNALAVLNDGDPINNTPTTGASPG 135
OY 155 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 214
DB 136 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 195
OY 215 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHEGCAAGCTGPKHSDDLAC 274
DB 196 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHEGCAAGCTGPKHSDDLAC 255
OY 275 LHFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVSGAGM 334
DB 256 LHFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVSGCT--L 313
OY 335 VHHRRSSSTRSGGD-----LTLGLEPSEEAAPRSPLAPSGAG----- 374
DB 314 VCPHLNOEVAEDGTORCEKSKPCARVCYGLGMEHLREVRAVTSANIOEFAGCKITFGS 373
OY 375 ----SDVFGD----- 381
DB 374 LAFLEPSFDDPASNTPALQPEQLQVETLEITGYLYISAMPDSDLDSVFONLQYIRG 433
OY 382 -----LGMGA--KGLOSL-----PTN---DPSPLQRYSEDP--- 408
DB 434 RILHNGAYSLTLOGIGISWLGSLRLSRELGSGLALIHNTHLCEFHVTYPMQDLFRNPQAL 493
OY 409 --TPVLPSEI--DGYVAPLTC-----PQ-----E 430
DB 494 LHFANRPEDCQVGEGLACHOLCARGHCGPGPTQCVNCSQFLRGQCEVECRVLQGLPRE 553
OY 431 YVN-----QPDVROPQPPRREGPLPAARPAATLERAKTLSPGKGVKDVAFGCAV 483
DB 554 YVANRHLCRPHREGCPQ-----NG-----SYTCFQPEA 581
OY 484 EN---PEVLTPOGGAAPORHPAPASPAFADNLYYWDODPERGA--PPSTF----- 529
DB 582 DQCVACAHYKDRPECYA---RCPSGVKRDLSYMPIM--KPFDEGACQPCPINCINHCVDL 637
OY 530 --RGTPTAENPEYLGIDVPAARARSPTSTQREHYNALQEARLLNLSRDTAAEMNET 587
DB 638 DDGCPRAEQNASPL---TSLAARARSPTSTQREHYNALQEARLLNLSRDTAAEMNET 694
OY 588 VEYISEAFDQEPCTCLQTRLELYKQGLRGSLTKLKGPLTMASHYKHOCPPETSCATO 647
DB 695 VEYISEAFDQEPCTCLQTRLELYKQGLRGSLTKLKGPLTMASHYKHOCPPETSCATO 754
OY 648 ITTFESFENKLDLFLVLPFCWEPVOE 675
DB 755 ITTFESFENKLDLFLVLPFCWEPVOE 782

```

RESULT 11
 US-09-344-195-4
 Sequence 4, Application us/09344195
 Patent No. 6210662
 GENERAL INFORMATION:
 APPLICANT: Laus, Reiner
 Ruegg, Curtis L.
 Wu, Hongyu
 TITLE OF INVENTION: Immunostimulatory Compositions
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Ave. Suite 250
 City: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306

COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/344,195
 FILING DATE: 24-Jun-1999
 CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/146,283
 FILING DATE: 03-SEPT-1998

ATTORNEY/AGENT INFORMATION:
 NAME: Judge, Linda R.
 REGISTRATION NUMBER: 42,702
 REFERENCE/DOCKET NUMBER: 7636-0010.21
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-0880
 TELEFAX: 650-324-0960

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 782 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: homo sapiens
 INDIVIDUAL ISOLATE: GM-CSF-Her-2 fusion protein: Fig. 8
 SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 US-09-344-195-4

Query Match 59.5%; Score 2224; DB 4; Length 782;
 Best Local Similarity 59.2%; Pred. No. 3.6e-165;
 Matches 478; Conservative 18; Mismatches 104; Indels 208; Gaps 20;

```

OY 35 LARGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVQVGLLETTYPTNASTLSFLQ 94
DB 16 LPRGAASVCTGCTGDMKRLPASPETHLDMRLHYOCGVVQVGLLETTYPTNASTLSFLQ 75
OY 95 DIOEVQGVYLIANNOVROVFLQRLRIYRGTOLEFEDNALAVLNDGDPINNTPTTGASPG 154
DB 76 DIOEVQGVYLIANNOVROVFLQRLRIYRGTOLEFEDNALAVLNDGDPINNTPTTGASPG 135
OY 155 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 214
DB 136 GLRELQRLSTLEIKGGVLIQIRNPOLCYODTIIMKDIFFHKNNOALATLIDTNSRACHPC 195
OY 215 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHEGCAAGCTGPKHSDDLAC 274
DB 196 SPCKGSRGCGESSEDCOSLTRFYCAGGACRGKGPLPTDCHEGCAAGCTGPKHSDDLAC 255
OY 275 LHFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVSGAGM 334
DB 256 LHFHNSGICELHCPALVTYNTDTFESMPNDEGRYTFGASCVTACPYNYLSTDVSGCT--L 313
OY 335 VHHRRSSSTRSGGD-----LTLGLEPSEEAAPRSPLAPSGAG----- 374
DB 314 VCPHLNOEVAEDGTORCEKSKPCARVCYGLGMEHLREVRAVTSANIOEFAGCKITFGS 373
OY 375 ----SDVFGD----- 381
DB 374 LAFLEPSFDDPASNTPALQPEQLQVETLEITGYLYISAMPDSDLDSVFONLQYIRG 433
OY 382 -----LGMGA--KGLOSL-----PTN---DPSPLQRYSEDP--- 408
DB 434 RILHNGAYSLTLOGIGISWLGSLRLSRELGSGLALIHNTHLCEFHVTYPMQDLFRNPQAL 493

```

QY 409 --TVLPSET---DGYVAPLTCS-----PQP-----E 430
DB 494 LHTARPEDECEGLACHQCARHCWGPGPTQCVCNCSQFLRGQECVECHVLOGLPRE 553
QY 431 YVN-----QPDVAPRPPSPREGPLPARPAGATLERAKTSLPGKNGVYKDVAFGGAV 483
DB 554 YVNARHCLPCHECCPQ-----NG----SVTCFGEBA 581
QY 484 EN-----PEYLPVGGGAAPQHPHPPAFSPAFDNLVYWDODPPERGA-PPSTF----- 529
DB 582 DOCAVACAHYKDPPEFVCA---RCPSGVKPDLSYMPITW-KPPDEGACQCPPICTHSCVDL 637
QY 530 --KGTPTAENPEYGLDLVPAAPARSPSPSTQPMENAVAIQEARRLNLSRDTAEMNET 587
DB 638 DDKGCPAGQBRASPL---TSLEAPARSPSPSTQPMENAVAIQEARRLNLSRDTAEMNET 694
QY 588 VEVISSEMDLOEPCTQLTRLELYKQGLRGLSLTKLKGPLTMASHYKQHCPPPTSCATQ 647
DB 695 VEVISSEMDLOEPCTQLTRLELYKQGLRGLSLTKLKGPLTMASHYKQHCPPPTSCATQ 754
QY 648 ITTFSEFENLKDPLVTPFCQWEPVOE 675
DB 755 ITTFSEFENLKDPLVTPFCQWEPVOE 782

RESULT 12

US-09-630-155-2
Sequence 2, Application US/09630155
Patent No. 6414130
GENERAL INFORMATION:
APPLICANT: Doherty, Joni Kristin and Gail M. Clinton
TITLE OF INVENTION: HER-2 BINDING ANTAGONISTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVIS WRIGHT TREMAINE LLP
STREET: 1501 Fourth Avenue, 2600 Century Square
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: PC compatible
OPERATING SYSTEM: Windows95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/630,155
FILING DATE: 16-Jan-2001
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Davison, Barry L.
REGISTRATION NUMBER: 47,309
REFERENCE/DOCKET NUMBER: 49321-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 628-7621
TELEFAX: 206 628-7699
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: polypeptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-630-155-2

Query Match 43.0%; Score 1607; DB 4; Length 419;
Best Local Similarity 77.2%; Pred.No. 2e-117;
Matches 318; Conservative 9; Mismatches 55; Indels 30; Gaps 7;
QY 35 LARGAASVQVCTGDMKRLRLPASPEETHDMLHLVYGGCVVGVGNLELYLPNASTSLQ 94
DB 16 LPPGAASVQVCTGDMKRLRLPASPEETHDMLHLVYGGCVVGVGNLELYLPNASTSLQ 75

QY 95 DIOEVGYVILAHNOVROVPIORLIRIVGTOLFEDNVALLAVLDNGDPLNNTTPVTGASPG 154
DB 76 DIOEVGYVILAHNOVROVPIORLIRIVGTOLFEDNVALLAVLDNGDPLNNTTPVTGASPG 135
QY 155 GLRELQLRSLTEILKGVLIQNRNPOLCYODTILMDIFHRKNQALTLTIDTRSRACHPC 214
DB 136 GLRELQLRSLTEILKGVLIQNRNPOLCYODTILMDIFHRKNQALTLTIDTRSRACHPC 195
QY 215 SPMCKGSRCKWGESSEDCSLRTVCAGCARKGFLPTDCCHEQCAAGCTGKRHSDCLAC 274
DB 196 SPMCKGSRCKWGESSEDCSLRTVCAGCARKGFLPTDCCHEQCAAGCTGKRHSDCLAC 255
QY 275 LHFNSGICELCPALVYNTDTPESMPNPEGRTYFGASCYACPYNTLSTVSGAGAG 334
DB 256 LHFNSGICELCPALVYNTDTPESMPNPEGRTYFGASCYACPYNTLSTVSGAGAG 313
QY 335 VHHRRSSSTRSGGDLTLGLEPSEEARPS--PLAPSBGAGSDVEDDLGMAAKGLQS 392
DB 314 VCPRLNQEVTAEADG---TQRCCKSKPCARGTHSLPRPA-----VPVPLRMQP 360
QY 393 LPTHDPSPLOKRSQD-----PTVPLPSETDGYVAPLTCSPOPEYVNOPD 436
DB 361 GPAPHVLSFLRPSWDLVSAFYSPLAPL-SPTSVSPISVSGRGPD-----PD 497

RESULT 13

US-08-422-108-1
Sequence 1, Application US/08422108
Patent No. 6015567
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Shepard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,108
FILING DATE: 14-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/355460
FILING DATE: 13-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-MAY-1989
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554CD22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid

TOPOLOGY: Linear
US-08-422-108-1

Query Match
Best Local Similarity 79.8%; Score 1597; DB 3; Length 624;
Matches 308; Conservative 6; Mismatches 34; Indels 38; Gaps 4;

QY 41 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 100
DB 1 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 60
QY 101 GYVLIANOVROVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTVTGASPGGLRELQ 160
DB 61 GYVLIANOVROVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTVTGASPGGLRELQ 120
QY 161 LRSITELIKGVLQIRNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPCKG 220
DB 121 LRSITELIKGVLQIRNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPCKG 180
QY 221 SRCGESSEDCQSLTRTVACAGGACRCKGRLPTDCHEQCAAGCTGPRHSDCLACLFHNS 280
DB 181 SRCGESSEDCQSLTRTVACAGGACRCKGRLPTDCHEQCAAGCTGPRHSDCLACLFHNS 240
QY 281 GICELHPALVTYNTDFEESMPNEGRTYGASCVTACPYNYLSTDVSGAGGVHHRHR 340
DB 241 GICELHPALVTYNTDFEESMPNEGRTYGASCVTACPYNYLSTDVSGAGGVHHRHR 298
QY 341 SSTRSGCD-----LTGLPSESEAPRSPPLASBGAG-----SD 376
DB 299 QEVTAEDGTQCEKCKSPCARVCYGLGMEHLREVRAVTSANIOGFACKKIFGSLAFLE 358
QY 377 VPDGDLGMAKGLQSLPTHDPSFLQ 402
DB 359 SFDGD-----PASNTAPLQ 372

RESULT 14
US-08-422-734-1
Sequence 1, Application US/08422734
Patent No. 6333169
GENERAL INFORMATION:
APPLICANT: Hudziak, Robert M.
APPLICANT: Sheppard, H. Michael
APPLICANT: Ullrich, Axel
TITLE OF INVENTION: HER2 EXTRACELLULAR DOMAIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,734
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422108
FILING DATE: 14-Apr-1995
APPLICATION NUMBER: 08/354460
FILING DATE: 13-Dec-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/048346
FILING DATE: 15-Apr-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/354319
FILING DATE: 19-May-1989

ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 554C2D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-422-734-1

Query Match
Best Local Similarity 79.8%; Score 1597; DB 4; Length 624;
Matches 308; Conservative 6; Mismatches 34; Indels 38; Gaps 4;

QY 41 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 100
DB 1 STGYCTGDMKRLRPASPEHLDMLRHLRYGCGVYVGNLETTYLPTNASLSFLDIOEVQ 60
QY 101 GYVLIANOVROVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTVTGASPGGLRELQ 160
DB 61 GYVLIANOVROVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTVTGASPGGLRELQ 120
QY 161 LRSITELIKGVLQIRNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPCKG 220
DB 121 LRSITELIKGVLQIRNPOLCYODTIIMKDIFFHKNOLATLTLIDTNSRACHPCSPCKG 180
QY 221 SRCGESSEDCQSLTRTVACAGGACRCKGRLPTDCHEQCAAGCTGPRHSDCLACLFHNS 280
DB 181 SRCGESSEDCQSLTRTVACAGGACRCKGRLPTDCHEQCAAGCTGPRHSDCLACLFHNS 240
QY 281 GICELHPALVTYNTDFEESMPNEGRTYGASCVTACPYNYLSTDVSGAGGVHHRHR 340
DB 241 GICELHPALVTYNTDFEESMPNEGRTYGASCVTACPYNYLSTDVSGAGGVHHRHR 298
QY 341 SSTRSGCD-----LTGLPSESEAPRSPPLASBGAG-----SD 376
DB 299 QEVTAEDGTQCEKCKSPCARVCYGLGMEHLREVRAVTSANIOGFACKKIFGSLAFLE 358
QY 377 VPDGDLGMAKGLQSLPTHDPSFLQ 402
DB 359 SFDGD-----PASNTAPLQ 372

RESULT 15
US-08-414-417B-69
Sequence 69, Application US/08414417B
Patent No. 5801005
GENERAL INFORMATION:
APPLICANT: Cheever, Martin A.
APPLICANT: Disis, Mary L.
TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed and Berry LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/414,417B


```

: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEO ID NO: 69:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 580 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
:
: US-08-414-417B-69

```

```

Query Match      31.6%; Score 1183; DB 1; Length 580;
Best Local Similarity 95.6%; Pred. No. 3.8e-84;
Matches 219; Conservative 0; Mismatches 4; Indels 6; Gaps 1;

```

```

QY 318 CPYVLTLDVSGAGGWHNRHRSSTRSGGDLTLGLEPSEERAPRSLAPSGAGSDV 377
Db 357 CP-----DPAAGAGGWHNRHRSSTRSGGDLTLGLEPSEERAPRSLAPSGAGSDV 410
QY 378 FDGDLGNGAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNOPDV 437
Db 411 FDGDLGNGAAGLQSLPTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNOPDV 470
QY 438 RPQPPSPREGPLPAARPAATLERAKTLSPGKNGVAKDVFAFGAVENPEYLTPQGAAP 497
Db 471 RPQPPSPREGPLPAARPAATLERAKTLSPGKNGVAKDVFAFGAVENPEYLTPQGAAP 530
QY 498 QPHPPAPSPAFDNLVYWDODPERGAPSPSTFKGPTAENPEYLGIDVP 546
Db 531 QPHPPAPSPAFDNLVYWDODPERGAPSPSTFKGPTAENPEYLGIDVP 579

```

Search completed: April 28, 2003, 13:43:11
 Job time : 34.1923.secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:49 ; Search time 46.1538 Seconds
(without alignments)
3080.404 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAAASLSIGFLF.....EPVQEGAPPPRAAHNNHHN 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_todent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2124.5	56.8	1259	6 018735	018735 canis famli
2	1609.5	43.0	419	4 09UK79	09UK79 homo sapien
3	969	25.9	367	11 08R2X1	08R2X1 mus musculu
4	906	24.2	165	4 014256	014256 homo sapien
5	900.5	24.1	412	4 08WYV0	08WYV0 homo sapien
6	754.5	20.2	527	13 090836	090836 gallus gall
7	723	19.3	149	6 09BG66	09BG66 oryctolagus
8	717.5	19.2	643	11 09ERV6	09ERV6 mus musculu
9	717.5	19.2	653	11 09WVF5	09WVF5 mus musculu
10	717.5	19.2	1210	11 09EP98	09EP98 mus musculu
11	714.5	19.1	1209	11 09QX70	09QX70 rattus norv
12	704.5	18.8	478	11 09ESES0	09ESES0 rattus norv
13	697	18.6	331	4 09BUD7	09BUD7 homo sapien
14	664.5	17.8	1165	13 09YH40	09YH40 xiphophorus
15	654	17.5	144	6 09GL44	09GL44 macaca mula
16	641.5	17.2	599	13 09PSH2	09PSH2 gallus gall

17	608	16.3	1328	13 P79754	P79754 fugu rubrip
18	545	14.6	144	6 09MYK4	09MYK4 ovis aries
19	543	14.5	1433	5 09B1H9	09B1H9 anophelis g
20	528	14.1	146	6 08WN17	08WN17 equus caball
21	515	13.8	152	6 09SL10	09SL10 equus caball
22	427	11.4	141	11 08VH40	08VH40 sigmodon th
23	407	10.9	1137	13 09W6F6	09W6F6 gallus gall
24	398	10.6	150	6 09BG64	09BG64 oryctolagus
25	395	10.6	151	6 09BG65	09BG65 oryctolagus
26	377	10.1	1368	5 023821	023821 caenorhabdi
27	362	9.7	138	11 099391	099391 marmota mon
28	341	9.1	366	5 026369	026369 schistosoma
29	341	9.1	1717	5 026566	026566 schistosoma
30	331	8.9	334	5 026567	026567 schistosoma
31	331	8.9	342	5 026568	026568 schistosoma
32	323.5	8.7	1193	5 09Y1X8	09Y1X8 ephydratia f
33	291	7.8	1472	5 09U5A8	09U5A8 bombyx mori
34	289	7.7	1671	5 09NJV5	09NJV5 bioplamarli
35	274	7.3	1749	5 08T0W6	08T0W6 echinococu
36	273.5	7.3	89	11 088459	088459 mus musculu
37	273.5	7.3	1418	13 08UW83	08UW83 paraliichth
38	272.5	7.3	1418	13 093457	093457 scophthalmu
39	270.5	7.2	2144	5 09VD94	09VD94 drosophila
40	263	7.0	1358	13 073798	073798 xenopus lae
41	259.5	6.9	1412	13 08UW84	08UW84 paraliichth
42	256.5	6.9	1362	13 09PV24	09PV24 xenopus lae
43	256	6.8	1369	13 08UW85	08UW85 paraliichth
44	247	6.6	1368	13 08UW85	08UW85 paraliichth
45	238	6.4	469	11 063721	063721 rattus norv

ALIGNMENTS

RESULT 1

ID 018735 PRELIMINARY: PRT: 1259 AA.
AC 018735;
DT 01-JAN-1998 (TREMBLrel. 05, Last Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Erdb-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_Taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RT "CDNA cloning of erdb-2 from canine mammary gland."
RL Submitted (OGF-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1;
DR HSBP; P11362; IEC.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_kinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Kinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.
DR ProDom: PD000001; Euk_kinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1259 AA; 13798 MW; E37364D49C4ACD46 CRC64;


```

Db 314 VCPHNOEVTAEDEG---TORCEKSKP-----CARGTHSLP 346
OY 395 THDPSPLQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVAPQ-----P 441
Db 347 ---PRP-----AAVPL-----PLRNQPGPAHVLISFLRPSMDLVSAFYSLPLAP 387
OY 442 PSPREGPL 449
Db 388 LSPTSVP 395

RESULT 3
OBR2X1 PRELIMINARY: PRT: 367 AA.
AC OBR2X1:
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 40.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
KW Hypothetical protein.
SQ SEQUENCE 367 AA; 40163 MW; 0BE03395F9E101B0 CRC64;

Query Match 25.9%; Score 969; DB 11; Length 367;
Best Local Similarity 80.0%; Pred. No. 3,1e-64;
Matches 180; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

OY 322 YLSTDVSGAGGVVHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGASDVFDGD 381
Db 142 FFSDDPALGTGSTAHRHRSSSARSAGGELTLGLEPSEEPSPPLASSEGASDVFDGD 201
OY 382 LGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVAPQ 441
Db 202 LAVGVTLGQSLSPHDSPLQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVAPQ 261
OY 442 PSPREGPLPARRPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYTLAPAGTASQPH 501
Db 262 PLTEGEGPPRPPRAGATLERKTLPSPKNGVYKDVFAFGAVENPEYTLAPAGTASQPH 321
OY 502 PPAESPAPFNDLYWDODPPERGAPSTFGKTPTAENPEYTLGLDVP 546
Db 322 SPASPAPFNDLYWDONSSQEGPPSTFGKTPTAENPEYTLGLDVP 366

RESULT 4
O14256 PRELIMINARY: PRT: 165 AA.
ID O14256
AC O14256:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE C-erb B2/neu protein (Fragment).
GN C-ERB B2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=86070181; Pubmed=2999974;
RA Consens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
RA Seeburg P.H., Liberman T.A., Schlessinger J., Francke U.,
RA Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor

```

```

RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=94000386; Pubmed=8104414;
RA Sakar F.H., Bail D.E., Li Y.W., Crisman J.D.;
RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)
RT gene.";
RL DNA Cell Biol. 12:611-615(1993).
DR EMBL; M95667; AAC37531.1; -.
FT NON_TER
SQ SEQUENCE 165 AA; 17327 MW; A0C113BA308BF46B CRC64;

Query Match 24.2%; Score 906; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 5.4e-60;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 373 AGSDVFDGDLGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYV 432
Db 1 AGSDVFDGDLGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYV 60
OY 433 NOPDVPPQPSRPGPLPARRPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYTLTPQ 492
Db 61 NOPDVPPQPSRPGPLPARRPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYTLTPQ 120
OY 493 GGAAPQHPPPAPFNDLYWDODPPERGAPSTFGKTPTAEN 537
Db 121 GGAAPQHPPPAPFNDLYWDODPPERGAPSTFGKTPTAEN 165

RESULT 5
OBR2X1 PRELIMINARY: PRT: 412 AA.
ID OBR2X1
AC OBR2X1:
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 44.7 kDa protein.
GN PF3659.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.Q., Qin W.X., Zhao X.T.,
RA Wan D.F., Gu J.R.;
RT "Novel human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF318349; AL55856.1; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF02757; YLP; 2.
DR Prodom: PD000001; Euk_Pkinase; 1.
DR SMART; SK00219; TykC; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein.
SQ SEQUENCE 412 AA; 44702 MW; 034397FF3F27D2BC CRC64;

Query Match 24.1%; Score 900.5; DB 4; Length 412;
Best Local Similarity 75.3%; Pred. No. 4.7e-59;
Matches 183; Conservative 6; Mismatches 30; Indels 24; Gaps 5;

OY 322 YLSTDVSGAGGVVHRRSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGASDVFDGD 381
Db 142 FFSDDPALGTGSTAHRHRSSSSTRSGGDLTLGLEPSEEAAPRSPPLASSEGASDVFDGD 201
OY 382 LGMAAKGLQSLPTHDPSPLOQRYSEDPVPLPSETDGVAPLTCSPQPEYVNOQDVAPQ 441

```

```

DB 202 LGGAAGKGLQSLPTHDSPLOQRSEDPYVPLPSETDGYVAPLCSQPQEVVNPDPVRPOP 261
OY 442 PSBREGPLPAPAPAGATLEAKTILSPKNGVVDVFAFGAVENPELYTPQGAAPQPH 501
DB 262 PSBREGPLPAPAPAGATLEAKTILSPKNGVVDVFAFGAVENPELYTPQGAALS --- 318
OY 502 PPAFSAFDNLVYWDPPERGAPST--FKGTPYAEINPEYLGL-DVPAAPAPRSPSPST 558
DB 319 PTLILPS-----AQPSTSYTGRTHTQSG---GLHPAPSKGHLRORTOST 360
OY 559 QPW 561
DB 361 WVV 363

```

RESULT 6
090836 PRELIMINARY: PRT: 527 AA.

```

AC 090836;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE EGF/EGF-alpha receptor precursor.
GN C-ERBB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-92123214; PubMed-1732751;
RA Flickinger T.W., Mahle N.J., Kung H.-J.;
RT "An alternatively processed mRNA from the avian c-erbB gene encodes a
RT soluble, truncated form of the receptor that can block ligand-
RT dependent transformatio".
RL Mol. Cell. Biol. 12:883-893(1992).
DR EMBL: M77637; AAA48759.1;
DR InterPro: IPR000494; EGFRL_domain.
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR SMART: SM00261; FU; 2.
DR pfam: PF01030; Recep_L-domain; 2.
KW Receptor; Signal.
FT SIGNAL 1 28
FT CHAIN 29 527 EGF/EGF-ALPHA RECEPTOR.
FT SEQUENCE 527 AA; 58353 MW; 764564ABCC095298 CRC64;

```

Query Match 20.2%; Score 754.5; DB 13; Length 527;
Best Local Similarity 42.4%; Pred. No. 5e-48;
Matches 160; Conservative 56; Mismatches 126; Indels 35; Gaps 10;

```

OY 19 LELLFWLDRSVLAKELARAASVOYCTGDMRLRLPASETHLMDLRHLHYOCQVYQGN 78
DB 18 LVLLLLGLVALCS---AVEEKVQCQGTNNLTQLGHVEDHTSLQRYNNCEVVLN 73
OY 79 LELTYLPTNALSFLDIOEVGCVLIANNOVQVPLQRLRIYRQQLFEDNYALAVLDN 138
DB 74 LETTYVEHNDLFLFKTIOGVAVIALLMNVIPLENOIIRGNVLYNSRALAVLSN 133
OY 139 GDRPLNTPVTVGASPGGLARELRSLEILKGVLIQRPOLCYODTIIMKDIIFHKNOL 198
DB 134 YH-MNKTQ-----GLRELPMKRLSEILNGCVKISNNKLCILNMPTVLNNDIIDSRK- 183
OY 199 ALVLID-TNRSRACHPCSPCKSGSRMGESSEDCOSTRIVVCGCA-RKGGPLPDCCH 256
DB 184 PLTVLDFASLSSCPKCHPCJEDHCWAGEQNCOTLVYCAQCCSGRCRGKVPDSCCH 243
OY 257 EDCAGCTGPKHSDCLACLHFNNSGICELCPALVYNTDTEFSMPNPEGRYTFGASCVT 316
DB 244 NCAACACTGCRSDCLACRFRDADATCKDCPPLVLYLXNPTTYMDVNPBEKYSFGATCVR 303

```

```

OY 317 ACPTNLTSDVVGAGAGVHHRRSSSTRSGCDLTLGLEPSEAPRSLABEGAGSD 376
DB 304 ECPHNVVYVDHGSVCV-----RSCNTDT-----YEVENGVRK-CRKCDCGCSK 345
OY 377 VFDGDLGMAKAGLOSL 393
DB 346 VCNG-IGIGELKGIILSI 361

```

RESULT 7
09BG66 PRELIMINARY: PRT: 149 AA.

```

AC 09BG66;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Receptor tyrosine kinase ErbB2 (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Klonisch T., Wolf P., Hombach-Klonisch S., Vogt S., Kuechenhoff A.,
RA Tetenus F., Fischer B.;
RT "ErbB genes and epidermal growth factor- (EGF-) like ligands in the
RT per-implantation rabbit uterus and blastocyst".
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF333178; AKK14371.1;
DR InterPro: IPR002174; Furin-like.
DR Pfam: PF00757; Furin-like; 1.
DR SMART: SM00261; FU; 2.
KW Kinase.
FT NON_TER 1 1
FT NON_TER 149 149
FT SEQUENCE 149 AA; 16240 MW; 7CB3792A54FC49BA CRC64;

```

Query Match 19.3%; Score 723; DB 6; Length 149;
Best Local Similarity 81.2%; Pred. No. 2.1e-46;
Matches 121; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

```

OY 177 NPQLCYODTLKMDIFKKNQALTLIDTNRSRACHPCSPCKSGSRMGESSEDCOSLTR 236
DB 1 NPQCYODTLQWQEFSTRTSWPSRTRINASRRATCPSPACAGCGSGSPEDCSLTR 60
OY 237 TVACGCGARCKGGLPTDCHEQCAAGCTGPKHSDCLACLHFNNSGICELCPALVYNTD 236
DB 61 TICAGCGARCKGGLPTDCHEQCAAGCTGPKHSDCLACLHFNNSGICELCPALVYNTD 120
OY 297 TFESMPNPEGRYTFGASCVTACPYNYLST 325
DB 121 TFESMPNPEGRYTFGASCVTTCPYNYLST 149
OY 297 TFESMPNPEGRYTFGASCVTACPYNYLST 325
DB 121 TFESMPNPEGRYTFGASCVTTCPYNYLST 149

```

RESULT 8
09ERV6 PRELIMINARY: PRT: 643 AA.

```

AC 09ERV6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Epidermal growth factor receptor isoform 2.
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-C3H/101, 129/SVJ, AND 129/SVEVYAC;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Sinclair C.S., Pearisall R.S., Green P.J., Yee D., Lampand A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,

```

RA Maible N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egr transcripts encoding truncated receptor
 RT isoforms";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF275366; AAG28046.1;
 DR EMBL: AF275364; AAG28046.1; JOINED.
 DR EMBL: AF275365; AAG28046.1; JOINED.
 DR MGI: 95294; Egr.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FU; 4.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR Receptor.
 KW SEQUENCE 643 AA; 71476 MW; DEF22002C84911BI CRC64;
 Query Match 19.2%; Score 717.5; DB 11; Length 643;
 Best Local Similarity 46.8%; Pred. No. 3.8e-45;
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;
 QY 36 ARGAA--STOVCTGDMKRLPASPETHLDMRLHYOCQVVOGNELELYPTNASISFL 93
 DB 20 AAGGALKEKVKCGTSNRLTOLGTFEDHFLSLQRMYNCEVGLNLEITYVQRYNDLSFL 79
 QY 94 ODIOEVOCYVLIHANOVQVPLQRLRIYRGQLPEFDNALVLDGDPDLNTPPTGASP 153
 DB 80 KTIQEVAGYVLIALTVPRIPLNQLITIGNALVNTALALISN-----YGTNR 129
 QY 154 GLRELQRLSTLEILKGVLIQNRNOLCYOPTIILMKDI-----FHNNOALATLIDTNSR 209
 DB 130 TGLRELPRNRNLQELILIGAVRSNNPILCNMDTIQRDIYVNVFMSMDL-----QSHPS 185
 QY 210 ACHPSPCKSKRCWGESSEDCQSLTRTVCAAGCA-RCKGPLPTDCHEGCAAGCTGPKH 268
 DB 186 SCPKCDPSCPMSGCGGEEHCQKLTITICAOQCSHRCGRSPSCCHNCAAGCTGPRE 245
 QY 269 SDCLACLFHNSGICELCPALVYNTDTPESMPREGRYFGASCYACPNYISTVG 328
 DB 246 SDCLVQCFQDEATCKDTPPLMLNPTTYQMDVNPESKYSFGAVCKCPRNYYVTDHG 305
 QY 329 S 329
 DB 306 S 306
 RESULT 9
 Q9WVF5 PRELIMINARY; PRT: 655 AA.
 AC Q9WVF5;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Epidermal growth factor receptor (Epidermal growth factor receptor
 DE isoform 3).
 DE EGFR.
 GN Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=LIVER.
 RA Røtter J.L., Threadgill D.W., Danielson A.J., Schell C.,
 RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
 RA Maible N.J.;
 RT "Alternative transcripts from the Human and Mouse EGFR Genes Encode
 RT Carboxy-Terminal Truncated Receptors";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVETAC;
 RA Røtter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
 RA Sinclair C.S., Peatsall R.S., Green P.J., Yee D., Lampland A.L.,
 RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
 RA Maible N.J.;
 RT "Comparative genomic sequence analysis and isolation of human and
 RT mouse alternative Egr transcripts encoding truncated receptor
 RT isoforms";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=LIVER;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi Y., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Kono H., Baldarelli R., Barish G.,
 RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta Y., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welt C., Whittaker C., Willing L.,
 RA Wyshew-Borls A., Yoshida K., Hasegawa Y., Kawai J., Kotsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL: AF124513; AAD44149.1;
 DR EMBL: AF275366; AAG28047.1;
 DR EMBL: AF275364; AAG28047.1;
 DR EMBL: AF275365; AAG28047.1; JOINED.
 DR EMBL: AK004944; BAB23688.1;
 DR EMBL: AK004883; BAB23641.1;
 DR EMBL: AK004911; BAB23662.1;
 DR MGI: 95294; Egr.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FU; 3.
 DR Receptor.
 KW SEQUENCE 655 AA; 72906 MW; 6B34063B1BC928CB CRC64;
 Query Match 19.2%; Score 717.5; DB 11; Length 655;
 Best Local Similarity 46.8%; Pred. No. 3.9e-45;
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;
 QY 36 ARGAA--STOVCTGDMKRLPASPETHLDMRLHYOCQVVOGNELELYPTNASISFL 93
 DB 20 AAGGALKEKVKCGTSNRLTOLGTFEDHFLSLQRMYNCEVGLNLEITYVQRYNDLSFL 79
 QY 94 ODIOEVOCYVLIHANOVQVPLQRLRIYRGQLPEFDNALVLDGDPDLNTPPTGASP 153
 DB 80 KTIQEVAGYVLIALTVPRIPLNQLITIGNALVNTALALISN-----YGTNR 129
 QY 154 GLRELQRLSTLEILKGVLIQNRNOLCYOPTIILMKDI-----FHNNOALATLIDTNSR 209
 DB 130 TGLRELPRNRNLQELILIGAVRSNNPILCNMDTIQRDIYVNVFMSMDL-----QSHPS 185
 QY 210 ACHPSPCKSKRCWGESSEDCQSLTRTVCAAGCA-RCKGPLPTDCHEGCAAGCTGPKH 268
 DB 186 SCPKCDPSCPMSGCGGEEHCQKLTITICAOQCSHRCGRSPSCCHNCAAGCTGPRE 245
 QY 269 SDCLACLFHNSGICELCPALVYNTDTPESMPREGRYFGASCYACPNYISTVG 328
 DB 246 SDCLVQCFQDEATCKDTPPLMLNPTTYQMDVNPESKYSFGAVCKCPRNYYVTDHG 305

Query Match 19.1%; Score 714.5; DB 11; Length 1209;
 Best Local Similarity 46.7%; Pred. No. 1.5e-44;
 Matches 142; Conservative 42; Mismatches 105; Indels 15; Gaps 5;

QY 30 VLAKELARGAA--STOVCTGTDMLRLPASPETHLMLRLHYOGCQVVOGNELETLPTN 87
 DB 14 LLAALCAAGALEEKVKVQCGTSNRLTQGFEDHFLSLQRMFNNECEVVLGNLEITYQVRN 73
 QY 88 ASLSFLDDIOEVGYVLIANNQVAPLOLRIRVGTQLEEDNALVALNDGPDLANNTTP 147
 DB 74 YDSFLTTIOEVAGVYLIANTVERIPLENQIIRGNALYENTYALVLSN----- 124
 QY 148 VTGASPGLELRLSLTEILKGVLIQRPOLCYODTILMKDIFHKNMOLATLIDTNR 207
 DB 125 -YGNKNGLELRLPRLNDELIGAVRESNPNILCMETIQMDIV-QDVFLSNMSMDVQR 182
 QY 208 S-RACHPCSPMCKGRCWGSESDCQSLFTVCAGGA-RCKGPLPDCHEGCAAGCTG 265
 DB 183 HLTGCPKDCSPMGSCMGEECQKLTIKCAQCCSRRCGRSPDCCHNCAAGCTG 242
 QY 266 PKHSDCLACHFNHSGICELHCPALVYNTDFESMNPREGRYFGASCYACRYNLTST 325
 DB 243 PRSDCLVCHRFDEATKDTCPPLMLNPTTYQMDVNPBGKYSFGATCVAKCPRYNVT 302
 QY 326 DVGS 329
 DB 303 DHGS 306

RESULT 12
 Q9ESED0 PRELIMINARY; PRT; 478 AA.
 ID Q9ESED0
 AC Q9ESED0
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Epidermal growth factor receptor related protein.
 GN ERFP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCB1_TaxID-101116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-GASTRO-DUODENAL MUCOSA;
 RA Yu Y., Mosher J.A., Majumdar A.P.N.;
 RT Cloning of a novel EGFR-related peptide: A putative negative
 RT regulator of EGFR.
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF187818; AAC17037.2; -
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like.
 DR Pfam: PF01030; Recep_L.domain; 2.
 KW Receptor.
 SQ SEQUENCE 478 AA; 53233 MW; CF873A8376G519E5 CRC64;

Query Match 18.8%; Score 704.5; DB 11; Length 478;
 Best Local Similarity 34.5%; Pred. No. 2.4e-44;
 Matches 180; Conservative 62; Mismatches 180; Indels 99; Gaps 18;

QY 36 ARGAA--STOVCTGTDMLRLPASPETHLMLRLHYOGCQVVOGNELETLPTNASTFL 93
 DB 20 AAGALEEKVAVCGTSNRLTQGFEDHFLSLQRMFNNECEVVLGNLEITYQVRNYSFL 79
 QY 94 QDIOEVGYVLIANNQVAPLOLRIRVGTQLEEDNALVALNDGPDLANNTTPVTGASP 153
 DB 80 KTIQEVAGVYLIANTVERIPSEDLQIRGNALYENTYALVLSN-----YGTNR 129
 QY 154 GGLRELQRLSLTEILKGVLIQRPOLCYODTILMKDI---EKKNNQALFTLIDTNRSR 209
 DB 130 TGLRELPMRLQDELIGAVRESNPNILCMNDTIQMDIVQNVFMSNMSMDL---QSHPS 185

QY 210 ACHPCSPMCKGRCWGSESDCQSLFTVCAGGA-RCKGPLPDCHEGCAAGCTGPKH 268
 DB 186 SCPCDCSPMGSCWGGEENCKQKLTIKCAQCCSHRCGRSPDCCHNCAAGCTGPQK 245
 QY 269 SDCLACHFNHSGICELHCPALVYNTDFESMNPREGRYFGASCYACRYNLTSDVG 328
 DB 246 SDCLVQCKFQDEATKDTCPPLMLNPTTYQMDVNPBGKYSFGATCVAKCPRYNVT 305
 QY 329 SGAGGAWHHRRSSTSGGDDLTLGLEPSEEPKRSPLASBEAGSYVFGDGLMKAAK 388
 DB 306 -----SCVACGPD---YVEVEDIRK-CKKDCPCRRKVCNG-IGIGFK 346
 QY 389 GLQSLPTHDPSPLRYSEDPVPLPSETDGYVAPLTCSPQREYVNOQDVRPQPSREGP 448
 DB 347 DTLSI---NATNIKHFYCTAI---SGDLHLPLVA-----FKGDSFTRTPPLDPR 391
 QY 449 LPAARPAATLERAKTSLSPGKNGV-----KDVAFGAVENPEYL---TPQG-- 494
 DB 392 -----LEIKTYKEITGSLIIQAMPENMTDLNAF---ENLEIRGRTRKHGQFS 437
 QY 495 -----AAPRHPRAFS-----PAFDNLTYMDOP 519
 DB 438 LAVVGLNTSLPQVPSLSMQAVTRPLHLAQNRYVSDTGP 478

RESULT 13
 Q9BUD7 PRELIMINARY; PRT; 331 AA.
 ID Q9BUD7
 AC Q9BUD7
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to v-erb-b2 avian erythroblastic leukemia viral oncogene
 DE homolog 3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCB1_TaxID-9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-PLACENTA;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC002706; AAH02706.1; -
 DR InterPro: IPR000494; EGFR_L.domain.
 DR InterPro: IPR002174; Furin-like.
 DR Pfam: PF00757; Furin-like.
 DR Pfam: PF01030; Recep_L.domain; 1.
 DR SMART: SM00261; FU; 2.
 SQ SEQUENCE 331 AA; 36489 MW; 45B8EBE683PE7E8 CRC64;

Query Match 18.6%; Score 697; DB 4; Length 331;
 Best Local Similarity 44.6%; Pred. No. 5.3e-44;
 Matches 139; Conservative 45; Mismatches 98; Indels 30; Gaps 7;

QY 16 LGFLFLFEMWDRSVLAKELARGAA--STOVCTGTDMLRLPASPETHLMLRLHYOGC 72
 DB 10 LGFLF-----SLARGSEVNGSAVCGTGLNGLSVTGDAENQVOTLTKKLERC 56
 QY 73 QVVOGNELETLPTNASTFLQDIOEVGYVLIANNQVAPLOLRIRVGTQLEEDNYA 132
 DB 57 EVVAGNLEIYLTGNADLSFLQIMREYTGVLVAMNESTLPLRLNRYVRSQVYDGF 116
 QY 133 LAVLDNDGPLNNTTPVTGASPGLELRLSLTEILKGVLIQRPOLCYODTILMKDIF 192
 DB 117 IFVW-----LNYNP-----NSHALRQLRLTQLEITLSCGVYIEKNDKLCHMDTIDMDIY 167
 QY 193 HKNQALFTLIDTNRSRACHPCSPMCKGRCWGSESDCQSLFTVCAGGA-AACKGPLP 251
 DB 168 RDRD---AETIVKNGKSGPCPCHEVCKG-RCKWGSESCQKLTITTCAPQNGHCFGP 223
 QY 252 TDCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVYNTDFESMNPREGRYT 311

```

Db 224 NQCHBCACGCGSPDQTDCAFCHRFNDGACVPCPQPLVYKNTLFOLEPNPHTKYQY 283
QY 312 ASCVTACPYNYL 323
Db 284 GVCVASCPHNFV 295

RESULT 14
QYH40
ID 09YH40 PRELIMINARY: PRT: 1165 AA.
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Receptor tyrosine kinase proto-oncogene.
GN XMRK.
OS Xiphophorus xiphidium.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorphia; Atherinomorphia;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=8086;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION;
RX MEDLINE=98241172; PubMed=9582016;
RA Dimitrijevic N., Winkler C., Wellbrock C., Gomez A., Duschl J.,
RA Altschmidt J., Schartl M.;
RT "Activation of the Xmrk proto-oncogene of Xiphophorus by
RT overexpression and mutational alterations.";
RL Oncogene 16:1681-1690(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO PURIFICATION;
RA Schartl M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: U53471; AAD10500.2;
DR HSSP: P11362; 1FGK.
DR InterPro: IPR000345; Cyt_c_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_kinase; 1.
DR SMART: SM00261; FUF; 3.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50019; PROTEIN_KINASE_DOM; 2.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
KW SEQUENCE 1165 AA; 129614 MW; 7F7EB38D771A74E CRC64;

Query Match 17.8%; Score 664.5; DB 13; Length 1165;
Best Local Similarity 41.4%; Pred. No. 7.4e-41;
Matches 143; Conservative 46; Mismatches 129; Indels 27; Gaps 9;
QY 43 QVCTGDMKLRAPSPETHLDMRLHLYOGGVQVGNLELYLTPTNLSLFLQDIOEVQY 102
Db 28 KYCGTSGNMTM--LDNHYLKMKKMYSGCNVYLENLEITYTQENDLSLQSIQEVGY 84
QY 103 VLIANQVQVPLQRLRIYRGITOLFEDNYALAVLDGDPPLNTTPTGASPGGLRELQRL 162
Db 85 VLIAMNEVSTPLVNLRLINGONLYEGNFTLWMSYOK--DYYQVGLKQLOLS 141
QY 163 SLTEILKGVLLIQRNPQLCYQDTILMKDIFHKNNQLATLILIDNRSRACHPCSPMKGSR 222

```

```

Db 142 NLTEILSGGVKSVSHNPFLCAVETIMMWDIVDKTSNPTMNLIPHAFFEQCCRCPCGCVNGS 201
QY 223 CMGSESEDCOSLIRPTVAGGC-ARCKGPLRPLDCCHECCAGACGPKKSDCLACHFNHSG 281
Db 202 CMAPGPGCHCKFTKLLCAECCNRRCRGPRIIDCCNEHCAGCGCTPRATDCLACDFDNDG 261
QY 282 ICCLHCPALVTYNTDFESMPNPEGRYTFGASCVTACPYNYLSTDVSGAGVHHRRS 341
Db 262 TCKDTCPPKRYDIVSHQVVDNPNIKRYTFGACVACVKECPSNVYVTE-----GACV----RS 312
QY 342 SSTRSGGGLTLGLEPSEEARPSPLAPSGAGSDVVDGLMGCA 386
Db 313 CSA-----GMLEVDENGKRS-CKPCDGVCPKVCDC-IGIGS 346

RESULT 15
QYGL44
ID 09GL44 PRELIMINARY: PRT: 144 AA.
AC 09GL44;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor.
GN GM-CSF.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Hutchinson K.L., Villinger F., Miranda M.E., Ksiazek T.G.,
RA Peters C.J., Rollin P.E.;
RT "Multiplex analysis of cytokines in the sera of cynomolgus macaques
RT naturally infected with Ebola (Reston).";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF007376; AAG16626.1;
DR HSSP: P04141; 2GMF.
DR InterPro: IPR000773; GM-CSF.
DR Pfam: PF01109; GM-CSF; 1.
DR PRINTS: PR00693; GMCSFACTOR.
DR ProDom: PD007349; GM-CSF; 1.
DR SMART: SM00040; CSF2; 1.
DR PROSITE: PS00702; GM-CSF; 1.
FT VARIANT 60 60 V -> I.
SQ SEQUENCE 144 AA; 16177 MW; 7D5F381DA2FC832F CRC64;

Query Match 17.5%; Score 654; DB 6; Length 144;
Best Local Similarity 94.5%; Pred. No. 2.8e-41;
Matches 121; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
QY 548 AAPARSPSPSTQPMENHNAIOEARRLNLSRDTAAENKTEVEYISEMFDLQEPCTCLOTRL 607
Db 17 SARARSPSPSTQPMENHNAIOEARRLNLSRDTAAENKTEVEYISEMFDLQEPSCLOTRL 76
QY 608 ELYKQGLSGSLTYLKGPPLYTMASHYKQCHPPTETSCATQITTFESKKNLKPFLVYIP 667
Db 77 ELYKQGLSGSLTYLKGPPLYTMASHYKQCHPPTETSCATQITTFOSKKNLKPFLVYIP 136
QY 668 DCWEPEVOE 675
Db 137 DCWEPEVOE 144

```

Search completed: April 28, 2003, 13:41:31
Job time : 52.1538 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:23 ; Search time 14.4231 Seconds

(Without alignments)
1984.228 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739
Sequence: 1 MRAAPLLARASLSIGFLF.....EPVQEGAPPPAAHHHHH 630

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	1	ERB2_HUMAN
2	1947.5	52.1	1257	1	ERB2_RAT
3	1935	51.8	1254	1	ERB2_MESAU
4	755.5	20.2	703	1	EGFR_CHICK
5	721	19.3	1210	1	EGFR_HUMAN
6	721	19.3	1308	1	ERB4_HUMAN
7	717.5	19.2	1210	1	EGFR_MOUSE
8	717	19.2	1308	1	ERB4_RAT
9	697	18.6	1342	1	ERB3_HUMAN
10	674	18.0	144	1	CSF2_HUMAN
11	668	17.9	1339	1	ERB3_RAT
12	655.5	17.5	1167	1	XMRK_XIPMA
13	549	14.7	144	1	CSF2_SHEEP
14	536	14.3	1426	1	EGFR_DROME
15	529	14.1	144	1	CSF2_CEREL
16	488	13.1	144	1	CSF2_PIG
17	481.5	12.9	143	1	CSF2_BOVIN
18	476.5	12.7	144	1	CSF2_CANFA
19	452	12.1	144	1	CSF2_FELCA
20	441	11.8	127	1	CSF2_RAT
21	435.5	11.6	140	1	CSF2_CAVPO
22	404.5	10.8	1323	1	LT23_CAEEL
23	367	9.8	141	1	CSF2_MOUSE
24	317.5	8.5	1363	1	ILPR_BRALA
25	293	7.8	245	1	ERB2_MOUSE
26	273.5	7.3	1477	1	HTK7_HYDAT
27	270.5	7.2	2146	1	INSR_DROME
28	257.5	6.9	1382	1	INSR_HUMAN
29	252.5	6.8	1372	1	INSR_MOUSE
30	252	6.7	1607	1	MIPR_LYMS
31	251.5	6.7	1300	1	IRR_MOUSE
32	249	6.7	1383	1	INSR_RAT
33	248.5	6.6	1297	1	IRR_HUMAN

34	246.5	6.6	1300	1	IRR_CAVPO	P14617 cavia porce
35	246	6.6	1367	1	IGIR_HUMAN	P6063 homo sapien
36	235.5	6.3	1370	1	IGIR_RAT	P24063 rattus norv
37	233.5	6.2	1373	1	IGIR_MOUSE	O60755 mus musculu
38	228.5	6.1	581	1	IRR_RAT	O64718 rattus norv
39	207	5.5	1390	1	INSR_AEDAE	O93105 aedes aegypt
40	178	4.8	2715	1	TRX2_HUMAN	O93105 homo sapien
41	168	4.5	386	1	PPAP_HUMAN	P15305 homo sapien
42	166	4.4	707	1	SEPO_HUMAN	P23246 homo sapien
43	162	4.3	1321	1	IRRS2_MOUSE	P81122 mus musculu
44	160	4.3	830	1	SREC_HUMAN	O14162 homo sapien
45	157.5	4.2	605	1	BRL1_EBV	P03209 Epstein-Bar

ALIGNMENTS

RESULT 1	ID	ERB2_HUMAN	STANDARD	PRT: 1255 AA.
AC	P04626			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (tyrosine kinase-type cell			
DE	surface receptor HER2) (MLN 19).			
GN	ERBB2 OR HER2 OR NCL OR NEU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=66118663; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,			
RT	"Similarity of protein encoded by the human c-erbB-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86016729; PubMed=2999974;			
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,			
RA	Francine U., Levinson A., Ullrich A.;			
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
RN	[3]			
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=2999974;			
RA	Semba K., Kamata N., Toyoshima K., Yamamoto T.;			
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor receptor gene and is amplified in a			
RT	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
RN	[4]			
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehman A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	-1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,			
CC	ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A			
CC	POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-			
CC	ALPHA AND AMPHIREGULIN.			
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein			
CC	tyrosine phosphate.			
CC	-1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS			
CC	(POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			

CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
 CC RESIDUES (BY SIMILARITY).
 CC -1- POLYMORPHISM: THERE ARE FOUR ALLELES DUE TO THE VARIATIONS IN
 CC POSITIONS 654 AND 655. ALLELE B1 (654-ILE-ILE-655) HAS A FREQUENCY
 CC OF 0.782; ALLELE B2 (654-ILE-VAL-655) HAS A FREQUENCY OF 0.206;
 CC ALLELE B3 (654-VAL-VAL-655) HAS A FREQUENCY OF 0.012.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M11767: AAA35808.1: -.
 DR EMBL: M11761: AAA35808.1: JOINED.
 DR EMBL: M11762: AAA35808.1: JOINED.
 DR EMBL: M11763: AAA35808.1: JOINED.
 DR EMBL: M11764: AAA35808.1: JOINED.
 DR EMBL: M11765: AAA35808.1: JOINED.
 DR EMBL: M11766: AAA35808.1: JOINED.
 DR EMBL: M11730: AAA75493.1: -.
 DR EMBL: M12036: AAA35978.1: -.
 DR EMBL: X03363: CAA27060.1: -.
 DR PIR: A25491: A25491.
 DR PIR: A24571: A24571.
 DR HSSP: P11362: 1FGK.
 DR Gene: HGNC:3430; ERBB2.
 DR MIM: 164870; -.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR001245; Furin-like.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; kinase.1.
 DR Pfam: PF00757; Furin-like.1.
 DR Pfam: PF01030; Recep_L_domain.2.
 DR Pfam: PF02757; YLP.2.
 DR ProDom: PD000001; Euk_kinase.1.
 DR SMART: SM00261; FU.3.
 DR SMART: SM00219; TyrKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM.1.
 DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;
 KW transferase; tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KM Polymorphism.
 FT SIGNL 1 21
 FT CHAIN 22 1235
 FT DOMAIN 22 652
 FT TRANSMEM 653 675
 FT DOMAIN 676 1255
 FT DOMAIN 720 987
 FT NP_BIND 726 734
 FT BINDING 733 733
 FT ACT_SITE 845 845
 FT DISULFID 199 204
 FT DISULFID 199 212
 FT DISULFID 220 227
 FT DISULFID 224 235
 FT DISULFID 224 244
 FT DISULFID 236 252
 FT DISULFID 240 264
 FT DISULFID 255 264
 FT DISULFID 268 295
 FT DISULFID 299 311
 FT DISULFID 315 331
 FT DISULFID 334 338
 FT DISULFID 511 520
 FT DISULFID 515 528
 FT DISULFID 531 540
 FT DISULFID 544 560
 FT DISULFID 544 560

FT DISULFID 563 576
 FT DISULFID 567 584
 FT DISULFID 587 596
 FT DISULFID 600 623
 FT DISULFID 626 634
 FT DISULFID 630 642
 FT MOD_RES 1139 1139
 FT MOD_RES 1248 1248
 FT CARBOHYD 68 68
 FT CARBOHYD 124 124
 FT CARBOHYD 187 187
 FT CARBOHYD 259 259
 FT CARBOHYD 330 330
 FT CARBOHYD 571 571
 FT CARBOHYD 629 629
 FT VARIANT 654 654
 FT VARIANT 655 655
 FT VARIANT 655 655
 FT CONFLICT 1170 1170
 FT SEQUENCE 1255 AA; 137909 MM; 39E9DFDA04DCE962 CRC64;
 SQ
 Query Match 64.3%; Score 2405.5; DB 1; Length 1255;
 Best Local Similarity 41.1%; Pred. No. 2.2e-124;
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;
 QY 35 LARGASTQVCTGDMKRLPASPETHLDMRLHYOGCVVQGNLELYLPTNASLSFLQ 94
 DB 16 LPFGAASTQVCTGDMKRLPASPETHLDMRLHYOGCVVQGNLELYLPTNASLSFLQ 75
 QY 95 DIOEVQGYVLIANQVQVQLRLRYRGTOLEEDNATYALVLDNGDPLNNTPTTGASPG 154
 DB 76 DIOEVQGYVLIANQVQVQLRLRYRGTOLEEDNATYALVLDNGDPLNNTPTTGASPG 135
 QY 155 GLRELQRLSTLEIKGVLQORNPOLCYOPTILMKDIFHNNQALATLIDTNRSRACHPC 214
 DB 136 GLRELQRLSTLEIKGVLQORNPOLCYOPTILMKDIFHNNQALATLIDTNRSRACHPC 195
 QY 215 SPCKGSRGCGESEDQSLTRTYVACAGCARCKGRLPTDCHEQACAGCTGPKHSDCLAC 274
 DB 196 SPCKGSRGCGESEDQSLTRTYVACAGCARCKGRLPTDCHEQACAGCTGPKHSDCLAC 255
 QY 275 LHRHSGICELHCPALVTYNTDFEESMPNPEGRTFASCVTACPYVLTSDVGS----- 329
 DB 256 LHRHSGICELHCPALVTYNTDFEESMPNPEGRTFASCVTACPYVLTSDVGSCTLYVC 315
 QY 330 ----- 329
 DB 316 PLINQEVTAEDGTORCEKSKPCARVQYGLMEHLREAVRANTSANIOEFAGCKKIEGSLA 375
 QY 330 ----- 329
 DB 376 FLPESEFGDPASNTAPLPQEQLOVFETLEITGYLTISAMPDLPDLSEVQNLQVINGRI 435
 QY 330 ----- 329
 DB 436 LHNQAVSLTLQGLISWGLRSRELGSGLALHHTNTLCTVHTVPWDQLFRNPQALLH 495
 QY 330 ----- 329
 DB 496 TANRPEDECVEGLACHQLCARGHGKGPPTQVCNCSQFLRGQCEVCECHVLOGLPREYV 555
 QY 330 ----- 329
 DB 556 NARHCLPCHPEQOPQNSVTCFGEADQCAVAHYKDPFCVACRPSGVKPDLSYMPIWK 615
 QY 330 ----- 329
 DB 616 FPDEGACQPCPINCNSVLDKDCGPABORASPLTISAVVGLLVVGVFGIL 675
 QY 330 ----- 329
 DB 676 KRQQRKIRKTYMRRLDQETELVEPLTPSGAMPNQAQRIKETEELRKVKVVGSGAGCTYV 735

QY 330 ----- 329
 DB 736 KGIMPDGENKIPVAIKVIRENTSPKANKRILDEAVYMGVSGSPYVRLIGLITSTVQ 795
 QY 330 ----- 329
 DB 796 LVTQLMPYGCILDHVRENRGLSGQDLLNMCQJAKGMSYLEVDVLYHRLAARNLVKS 855
 QY 330 ----- 329
 DB 856 PNHVKTDFGLARLLIDETFEYHADGKVPKIMMALESILRRPTHOSDVMSYGVYWEI 915
 QY 330 ----- 329
 DB 916 MTFGAKPYDGIPIAREIDLEKGERLPQPICTIDVYIMYKCMWIDSECRPRELVE 975
 QY 330 ----- 329
 DB 976 FSRMARDFQRFVYIIONEDLGPASPLDSTFYRSLLEDDMDGLVDAEYLVPOGCFPCPD 1035
 QY 330 ----- 387
 DB 1036 APGAGGVHHRHRSSTRSGGDLTLGLEPSEEAAPRPLAPSGAGSDVFDGLGMAA 1095
 QY 388 KGLQSLPTHPDPSLPQKSEDPYVLPSPETDGYVAPLTCSPQEPYVNOPDVAPPPSPREG 447
 DB 1096 KGLQSLPTHPDPSLPQKSEDPYVLPSPETDGYVAPLTCSPQEPYVNOPDVAPPPSPREG 1155
 QY 448 PLPAPAPAGATLEBAKTLSPGKNGVAVDFAFGAVENPEVLTPOGGAAPRHPPPAPSP 507
 DB 1156 PLPAPAPAGATLEBAKTLSPGKNGVAVDFAFGAVENPEVLTPOGGAAPRHPPPAPSP 1215
 QY 508 AFDNLVYWDODPPRPGAPSPSTFKGTPTAENPEYGLDVP 546
 DB 1216 AFDNLVYWDODPPRPGAPSPSTFKGTPTAENPEYGLDVP 1254
 RESULT 2
 ERBB2_RAT STANDARD; PRT: 1257 AA.
 ID ERBB2_RAT
 AC P06494;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
 DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (epidermal growth factor receptor-related protein).
 GN ERBB2 OR NEU.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Neuroblastoma;
 RX MEDLINE=86118652; Pubmed=3945311;
 RA Bergmann C.I., Hung M.-C., Weinberg R.A.;
 RT "The neu oncogene encodes an epidermal growth factor receptor-related protein.";
 RL Nature 319:226-230(1986).
 RN [2]
 RP SEQUENCE OF 852-905 FROM N.A.
 RC TISSUE-Sciatic nerve;
 RX MEDLINE=91222560; Pubmed=2025425;
 RA Lai C., Lemke G.;
 RT "An extended family of protein-tyrosine kinase genes differentially expressed in the vertebrate nervous system.";
 RL Neuron 6:691-704(1991).
 RN [3]
 RP STRUCTURE BY NMR OF 650-668.
 RX MEDLINE=9215181; Pubmed=1346763;
 RA Gullick W.J., Bottomley A.C., Loftis F.J., Doak D.G., Mulvey D.,

RA Newman R., Crumpton M.J., Sternberg M.J.E., Campbell I.D.;
 RT "Three dimensional structure of the transmembrane region of the proto-oncogenic and oncogenic forms of the neu protein.";
 RL EMBO J. 11:43-48(1992).
 CC -1- FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX, ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-ALPHA AND AMPHIREGULIN.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein tyrosine phosphate.
 CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS.
 CC -1- THE CONSTITUTIVELY ACTIVATED ONCOGENIC VARIANT FORMS A HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X03362; CAA27059.1; ALT_INT.
 DR PIR: A24562; TYRPTNU.
 DR HSSP: P11362; IREG.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00068; Pkinase.1.
 DR Pfam: PF00757; Furin-like.1.
 DR Pfam: PF01030; Recep_L_domain.2.
 DR Pfam: PF02757; YLP.2.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR SMART: SM00261; Fu.3.
 DR SMART: SM00219; TyKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Proto-oncogene; Disease mutation.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1257 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT DOMAIN 22 654 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 655 677 POTENTIAL.
 FT DOMAIN 678 1257 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 159 369 CYS-RICH.
 FT DOMAIN 473 646 CYS-RICH.
 FT DOMAIN 722 989 PROTEIN KINASE.
 FT NP_BIND 728 736 ATP (BY SIMILARITY).
 FT BINDING 755 755 ATP (BY SIMILARITY).
 FT ACT_SITE 847 847 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 200 213 BY SIMILARITY.
 FT DISULFID 221 228 BY SIMILARITY.
 FT DISULFID 225 236 BY SIMILARITY.
 FT DISULFID 237 245 BY SIMILARITY.
 FT DISULFID 241 253 BY SIMILARITY.
 FT DISULFID 256 265 BY SIMILARITY.
 FT DISULFID 269 296 BY SIMILARITY.
 FT DISULFID 300 312 BY SIMILARITY.
 FT DISULFID 316 332 BY SIMILARITY.
 FT DISULFID 335 339 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 517 530 BY SIMILARITY.
 FT DISULFID 533 542 BY SIMILARITY.
 FT DISULFID 546 562 BY SIMILARITY.
 FT DISULFID 565 578 BY SIMILARITY.

```

FT DISULFID 569 586 BY SIMILARITY.
FT DISULFID 598 598 BY SIMILARITY.
FT DISULFID 602 625 BY SIMILARITY.
FT DISULFID 628 636 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT MOD_RES 1141 1141 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 1250 1250 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 631 631 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 661 661 V -> E (IN ONCOGENIC NEU).
SQ SEQUENCE 1257 AA; 138831 MW; 6129264583011402 CRC64;

Query Match 52.1%; Score 1947.5; DB 1; Length 1257;
Best Local Similarity 34.2%; Pred. No. 2.4e-99;
Matches 433; Conservative 21; Mismatches 70; Indels 741; Gaps 3;

OY 11 AASLSIGFLFLFWMIDRSYLAKELARGAASSTQVCTGDMKRLPASPETHLDMRLHYQ 70
DB 4 AAMCRMGFLAL-----LPPGIAGTQVCTGDMKRLPASPETHLDMRLHYQ 51
OY 71 GCOVVOGNLELYLPTNASTLFLQIDVOGYVLIANHOVROPLORLRIVGTQLPEDN 130
DB 52 GCOVVOGNLELYLPTNASTLFLQIDVOGYVLIANHOVROPLORLRIVGTQLPEDN 111
OY 131 YALAVDNDPLNNTTTPVT-GASPGGLRELOLSLREILKGVLIORNPOLCYOPTILMK 189
DB 112 YALAVDNDPLNNTTTPVT-GASPGGLRELOLSLREILKGVLIORNPOLCYOPTILMK 171
OY 190 DIFHKNNQALTLIDTNRSHACHPSCKSGRCMGESSEDCSLTRVYACAGCARCKGP 249
DB 172 DYFRKNNQALPVIDIDTNRSHACHPSCKSGRCMGESSEDCSLTRVYACAGCARCKGP 231
OY 250 LPTDCGHECCAACTGPKHSDCLACHFNHSGICELHCPALVYNNIDTPESMNPREGRT 309
DB 232 LPTDCGHECCAACTGPKHSDCLACHFNHSGICELHCPALVYNNIDTPESMNPREGRT 291
OY 310 FGASCTACPYNYLSTDVGS----- 329
DB 292 FGASCTACPYNYLSTDVGS----- 351
OY 330 ----- 329
DB 352 RGAARITSDNVQFEDGCKKIFGSLAFLEPSFDGDPSSGIALPRLPEQLQVETLEITGYL 411
OY 330 ----- 329
DB 412 YISAMPDSLRLDSVFQNLRIINGRIIHDGAYSLTLOGLIHSGLRLSRELGSGLALIR 471
OY 330 ----- 329
DB 472 NAHLCEVHYVPMQDLFRNPHQALLHSGNRPEEDLVSSGLVCSLCAHGHCMGPGLQCY 531
OY 330 ----- 329
DB 532 NCSHFLRGOECVECHVMKGLPREVYSDKRLCPHCECOPONSSETCFGEADCAACAH 591
OY 330 ----- 329
DB 592 YKDSSCVACRPSGVKPLSYPMWKYPDEEGICOPCPINCHSCVLDLDERGCPABORAS 651
OY 330 ----- 329
DB 652 PYFIIATVVGVLFLLLVYVVGILKRRKRIKRTMRLLDETLEVEPLPSPGAMPND 711
OY 330 ----- 329
DB 712 AQMRILKETELRKVKVLSGAGACTYVKGIMPDGENVKIPVAIKVLRNTPSKANKEILD 771
OY 330 ----- 329

```

```

DB 772 EAYVAVGSPYVSRLLGICITSTVOLVTLQMPYGCILLDHVREHGRGLSQDLLMNCVO1 831
OY 330 ----- 329
DB 832 AKCMSTLEDVRLVHRLAARNLVKSPNHVKTIDFGLARLLIDETERYHADGKVP1KMM 891
OY 330 ----- 329
DB 892 ALESLIRRTHQSDVMSYGVWELMTFGAKPYDGIPIAREIPDLLEKGERLPPICTI 951
OY 330 ----- 329
DB 952 DVMYIMKCMWIDSECRPRERELVSFRRMARDPORFVYIQNEDLGPSSPKSTFYRSL 1011
OY 330 ----- 361
DB 1012 EDDMDGLDYDAEYLYLPQGFSPDPTPGTGAHRRHSSSTRSGGELTLLEPSEEG 1071
OY 362 APRSPAPSEAGSDVFDGDLGMAKGLQSLPTPHDSPLOQRYSEDPVPLPSETDGYVA 421
DB 1072 PRSPPLAPSEAGSDVFDGDLGMAKGLQSLPTPHDSPLOQRYSEDPVPLPSETDGYVA 1131
OY 422 PLTCSQPEYVNOVDYRPPSPREGLPAARAGATLERAKTLSPGKNGVADVAFGC 481
DB 1132 PLTCSQPEYVNOVDYRPPSPREGLPAARAGATLERAKTLSPGKNGVADVAFGC 1191
OY 482 AVENPEYLPFGGAAPPPAFSPAFDNLVYWDOPPERGAPSTFGTFAENPEYL 541
DB 1192 AVENPEYLPREGTASPPHSPAFDNLVYWDONSSQGGPPNFECTFAENPEYL 1251
OY 542 GLDVP 546
DB 1252 GLDVP 1256

RESULT 3
ID ERB2_MESAU STANDARD: PRT: 1254 AA.
AC 060553;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2).
GN ERBB2 OR NEU.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Nerve;
RX MEDLINE=94193007; PubMed=7908275;
RA Nakamura T., Ushijima T., Ishizaka Y., Nagao M., Arai M.,
RA Yamazaki Y., Ishikawa T.;
RT *Cloning and activation of the Syrian hamster neu proto-oncogene.*;
RL Gene 140:251-255(1994).
CC - FUNCTION: ESSENTIAL COMPONENT OF A NEUREGULIN-RECEPTOR COMPLEX,
CC ALTHOUGH NEUREGULINS DO NOT INTERACT WITH IT ALONE. GP30 IS A
CC POTENTIAL LIGAND FOR THIS RECEPTOR. NOT ACTIVATED BY EGF, TGF-
CC ALPHA AND AMPHIREGULIN (BY SIMILARITY).
CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC - SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC - SUBCELLULAR LOCATION: Type I membrane protein.
CC - PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES.
CC - SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


QY 378 FDGDLGGAAGKGLQSLPTHPSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQDV 437
 DB 1086 FEGELGAGKARKGQSPISDRSLQRYSEDPPLPLPTETDGYVAPLTCSPQPEYVNPQDV 1145
 OY 438 RPOPSRRECPPLPAARAGATLERAKTLSPGKNGVADVAFGAVENPEYLPPOGGAAP 497
 DB 1146 RPOPSRRECPPLPAARAGATLERAKTLSPGKNGVADVAFGAVENPEYLPPOGGAAP 1205
 OY 498 QPHHPAFSPAFNLYVMDODPERGAPSPSTFKGTPTAENPEYLGDLVP 546
 DB 1206 QPHHPAFSPAFNLYVMDODPERGAPSPSTFKGTPTAENPEYLGDLVP 1253

RESULT 4
 EGF_R_CHICK STANDARD; PRT; 703 AA.
 ID EGF_R_CHICK
 AC P1387;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (CER)
 GN (fragment).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88261272; PubMed=3260329;
 RA Lax I., Johnson A., Howk R., Sap J., Bellot F., Winkler M.,
 RA Ullrich A., Vennstrom B., Schlessinger J., Givol D.;
 RT "Chicken epidermal growth factor (EGF) receptor: cDNA cloning,
 RT expression in mouse cells, and differential binding of EGF and
 RT transforming growth factor alpha".
 RL Mol. Cell. Biol. 8:1970-1978(1988).
 CC - FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 CC AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 CC VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC - CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC - SUBCELLULAR LOCATION: Type I membrane protein.
 CC - MISCELLANEOUS: Binding of EGF to the receptor leads to
 CC dimerization, internalization of the EGF-receptor complex,
 CC induction of the tyrosine kinase activity, stimulation of cell DNA
 CC synthesis, and cell proliferation.
 CC - SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M20386; AAA48760.1; -
 DR InterPro: IPR000494; EGF_R_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR SMART: SM00261; FUr; 4.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYP; PARTIAL.
 DR PROSITE: PS00107; PROTEIN_KINASE_TYP; PARTIAL.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; PARTIAL.
 KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
 KW Tyrosine-protein kinase; ATP-binding; Phosphorylation.
 FT SIGNAL 1 30
 FT CHAIN 31 >703
 FT DOMAIN 31 654
 FT EPIDERMAL GROWTH FACTOR RECEPTOR.
 FT EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 655 667
 FT DOMAIN 668 >703
 FT DISULFID 197 206
 FT DISULFID 201 214
 FT DISULFID 222 230
 FT DISULFID 226 238
 FT DISULFID 239 247
 FT DISULFID 243 255
 FT DISULFID 258 267
 FT DISULFID 271 298
 FT DISULFID 302 314
 FT DISULFID 318 333
 FT DISULFID 336 340
 FT DISULFID 513 522
 FT DISULFID 517 530
 FT DISULFID 533 542
 FT DISULFID 546 562
 FT DISULFID 565 581
 FT DISULFID 569 589
 FT DISULFID 592 601
 FT DISULFID 605 627
 FT DISULFID 630 638
 FT DISULFID 634 646
 FT CARBOHYD 134 134
 FT CARBOHYD 190 190
 FT CARBOHYD 200 200
 FT CARBOHYD 359 359
 FT CARBOHYD 368 368
 FT CARBOHYD 420 420
 FT CARBOHYD 573 573
 FT CARBOHYD 578 578
 FT CARBOHYD 613 613
 FT CARBOHYD 633 633
 FT CARBOHYD 648 648
 FT NON TER 703 703

SO SEQUENCE 703 AA; 77427 MW; AFD2E1B735A690 CRC64;

Query Match 20.2%; Score 755.5; DB 1; Length 703;
 Best Local Similarity 41.9%; Pred. No. 1,7e-34;
 Matches 161; Conservative 57; Mismatches 119; Indels 47; Gaps 10;

OY 28 RSVLAKELARGA-----STQVTCGDMKRLRASPETHLMDRLHYOG 71
 DB 4 RSVLAKELARGA-----STQVTCGDMKRLRASPETHLMDRLHYOG 71
 OY 72 COVQNGMLETLYPTNASTSLFDIOEVQGYVLIAHQVROVPLRLIYRGTOLEFDNY 131
 DB 64 CEVLSNLEITTYEHNNDLTFKTIQEVAGYVLLALMNVPIPLENQTIRGNVLYDINSF 123
 OY 132 ALAVLNGDPLNMTPTVYGASPGGLRELRLSLTEILKGVLLQIRNPOLCYODTILMKDI 191
 DB 124 ALAVLSNYH-MNKTQ-----GLRELPMKRSLSEILNGGVKISNNKLCNMIDVLMNDI 174
 OY 192 FHKRNQALFLID-TNRSRACHPCSPMKGRKSGESEDCCQSLTRVCAAGCA-RCKGP 249
 DB 175 IDTSRK-PLTVLPAFASLSSCPKCHPNCETEDHCGAGEQNCQFTTYKYLQAQCSGRGK 233
 OY 250 LPTDCHEQCAAGCTGKSHDCLALFHNHSGICELCPALVTYNTDFCSMPNPEGRYT 309
 DB 234 VPSSCCNQCAGCTGKSHDCLALFHNHSGICELCPALVTYNTDFCSMPNPEGRYT 293
 OY 310 FGASCVTACPYNYLSTDVSGAGAGGVHHRSSSTRSGGDDTLGLEPSEEARPLAP 369
 DB 294 FGATCVRECPHNYVTVTHGSCV-----RSCNTDT-----YVEENGVRK-CKK 335
 OY 370 SEGAGSDYFDGDLGMAKGLQSL 393
 DB 336 CDGLSKVNCG-IGIGELKGLIST 358

RESULT 5
 EGF_R_HUMAN STANDARD; PRT; 1210 AA.
 ID EGF_R_HUMAN

AC P00533; P06268; Q14225; Q9UMD7; Q9UMD8; Q9UMG5; Q92795; Q00732;
AC O00688; Q9B252; Q9H2C9; Q9G2X1; Q9H3C9;
DT 21-JUN-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Epidermal growth factor receptor precursor (EC 2.7.1.112) (Receptor
DE protein-tyrosine kinase ErbB-1).
GN EGFR OR ERBB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=84219729; PubMed=6328312;
RA Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A., Tam A.W.,
RA Lee J., Yarden Y., Libermann T.A., Schlessinger J., Downward J.,
RA Mayes E.L.V., Whittle N., Waterfield M.D., Seeburg P.H.;
RT "Human epidermal growth factor receptor cDNA sequence and aberrant
RT expression of the amplified gene in A431 epidermoid carcinoma cells.";
RL Nature 309:418-425(1984).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Placenta;
RC MEDLINE=95382957; PubMed=7654368;
RA Ilekis J.V., Stark B.C., Scoccia B.;
RT "Possible role of variant RNA transcripts in the regulation of
RT epidermal growth factor receptor expression in human placenta.";
RL Mol. Reprod. Dev. 41:149-156(1995).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Placenta;
RC MEDLINE=97078686; PubMed=8918811;
RA Reltter J.L., Maible N.J.;
RT "A 1.8 kb alternative transcript from the human epidermal growth
RT factor receptor gene encodes a truncated form of the receptor.";
RL Nucleic Acids Res. 24:4050-4056(1996).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Placenta;
RC MEDLINE=97256547; PubMed=9103388;
RA Ilekis J.V., Garfili J., Niederberger C., Scoccia B.;
RT "Expression of a truncated epidermal growth factor receptor-like
RT protein (TEGFR) in ovarian cancer.";
RL Gynecol. Oncol. 65:36-41(1997).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RX TISSUE=Placenta;
RC MEDLINE=21100872; PubMed=11161793;
RA Reltter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielsen A.J.,
RA Schehl Sinclair C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.D., Magnuson T.R., James C.D.,
RA Maible N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative EGFR transcripts encoding truncated receptor
RT isoforms.";
RL Genomics 71:1-20(2001).
RN [6]
RP SEQUENCE OF 575-687 FROM N.A.
RA Reltter J.L., Threadgill D.W., Danielsen A.J., Schehl C.M.,
RA Lampland A.L., Balasubramaniam S., Crossley T.O., Magnuson T.R.,
RA Maible N.J.;
RT "Human and mouse alternative EGFR transcripts encoding only the
RT extracellular domain of the receptor.";
RL submitted (Feb-1999) to the EMBL/Genbank/DBJ databases.
RN [7]
RP SEQUENCE OF 713-924 FROM N.A.
RX MEDLINE=84196372; PubMed=6326261;
RA Lin C.R., Chen W.S., Krulger W., Stolarsky L.S., Weber W.,
RA Evans R.M., Verna I.M., Gill G.N., Rosenfeld M.G.;
RT "Expression cloning of human EGF receptor complementary DNA: gene
RT amplification and three related messenger RNA products in A431
RT cells.";
RN Science 224:843-848(1984).
RN [8]
RP SEQUENCE OF 150-962 FROM N.A.
RX MEDLINE=84245835; PubMed=6330563;
RA Xu Y.H., Ishii S., Clark A.J.L., Sullivan M., Wilson R.K., Ma D.P.,
RA Roe B.A., Merlino G.T., Pastan I.;
RT "Human epidermal growth factor receptor cDNA is homologous to a
RT variety of RNAs overproduced in A431 carcinoma cells.";
RL Nature 309:806-810(1984).
RN [9]
RP SEQUENCE OF 1028-1210 FROM N.A.
RX MEDLINE=85046483; PubMed=6093780;
RA Slimen F.A., Gope M.L., Schultz T.Z., Wright D.A., Carpenter G.,
RA O'Malley B.W.;
RT "Isolation of an evolutionarily conserved epidermal growth factor
RT receptor cDNA from human A431 carcinoma cells.";
RL Biochem. Biophys. Res. Commun. 124:125-132(1984).
RN [10]
RP SEQUENCE OF 1-29 FROM N.A.
RX MEDLINE=88217333; PubMed=3329716;
RA Haley J.D., Whittle N., Bennett P., Kinchington D., Ullrich A.,
RA Waterfield M.D.;
RT "The human EGF receptor gene: structure of the 110 kb locus and
RT identification of sequences regulating its transcription.";
RN Oncogene Res. 1:375-396(1987).
RN [11]
RP SEQUENCE OF 1-29 FROM N.A.
RX MEDLINE=91107677; PubMed=1988448;
RA Haley J.D., Waterfield M.D.;
RT "Contributory effects of de novo transcription and premature
RT transcript termination in the regulation of human epidermal growth
RT factor receptor proto-oncogene RNA synthesis.";
RL J. Biol. Chem. 266:1746-1753(1991).
RN [12]
RP SEQUENCE OF 1-29 FROM N.A.
RX MEDLINE=85270438; PubMed=2991899;
RA Ishii S., Xu Y.H., Strutton R.H., Roe B.A., Merlino G.T., Pastan I.;
RT "Characterization and sequence of the promoter region of the human
RT epidermal growth factor receptor gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4920-4924(1985).
RN [13]
RP SEQUENCE OF 540.
RA Kohda D.;
RT submitted (SEP-1997) to the SWISS-PROT data bank.
RN [14]
RP RECEPTOR ACTIVITY.
RX MEDLINE=84191554; PubMed=6325948;
RA Mroczkowski B., Mosig G., Cohen S.;
RT "ATP-stimulated interaction between epidermal growth factor receptor
RT and supercoiled DNA.";
RL Nature 309:270-273(1984).
RN [15]
RP PHOSPHORYLATION.
RX MEDLINE=89278137; PubMed=2543678;
RA Margolis B.L., Lax I., Kris R., Dombalegian M., Honegger A.M.,
RA Hawk R., Givol D., Ullrich A., Schlessinger J.;
RT "All autophosphorylation sites of epidermal growth factor (EGF)
RT receptor and HER2/neu are located in their carboxyl-terminal tails.
RT Identification of a novel site in EGF receptor.";
RL J. Biol. Chem. 264:10667-10671(1989).
RN [16]
RP CARBOHYDRATE-LINKAGE SITES ASN-128; ASN-175; ASN-413; ASN-444 AND
RP ASN-528.
RX MEDLINE=86398132; PubMed=8962717;
RA Smith K.D., Davies M.J., Bailey D., Renouf D.V., Hounsell E.F.;
RT "Analysis of the glycosylation patterns of the extracellular domain of
RT the epidermal growth factor receptor expressed in Chinese hamster
RT ovary fibroblasts.";
RL Growth Factors 13:121-132(1996).
RN [17]
RP CARBOHYDRATE-LINKAGE SITES ASN-56; ASN-352; ASN-361; ASN-568 AND
RP ASN-603.
RX MEDLINE=20198209; PubMed=10731668;

RA	Sato C, Kim J.-H., Abe Y., Saito K., Yokoyama S., Kohda D.;
RT	"Characterization of the N-oligosaccharides attached to the atypical
RT	Asn-x-Cys sequence of recombinant human epidermal growth factor
RT	receptor.";
RL	J. Biochem. 127:65-72(2000).
RN	[18]
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX	MEDLINE-98225196; Pubmed-9556602;
RA	Abe Y., Ooka M., Inagaki F., Iax I., Schlessinger J., Kohda D.;
RT	"Disulfide bond structure of human epidermal growth factor receptor."
RT	J. Biol. Chem. 273:11150-11157(1998).
RL	[19]
RN	REVIEW.
RX	MEDLINE-87297456; Pubmed-3039909;
RA	Carpenter G.;
RT	"Receptors for epidermal growth factor and other polypeptide
RT	mitogens.";
RL	Annu. Rev. Biochem. 56:681-914(1987).
CC	-I- FUNCTION: Receptor for EGF, but also for other members of the EGF
CC	family, as TGF-alpha, amphiregulin, betacellulin, heparin-binding
CC	EGF-like growth factor, GP30 and vaccinia virus growth factor. Is
CC	involved in the control of cell growth and differentiation.
CC	-I- FUNCTION: Activity 2/truncated isoform may act as an antagonist.
CC	-I- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC	tyrosine phosphate.
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein. Isoform 2 is
CC	secreted.
CC	-I- ALTERNATIVE PRODUCTS: 4 isoforms; 1/p170 (shown here), 2/p60/
CC	truncated isoform/TEGFR, 3/p110 and 4; are produced by
CC	alternative splicing.
CC	-I- TISSUE SPECIFICITY: Expressed in placenta. Isoform 2 is also
CC	expressed in ovarian cancers.
CC	-I- MISCELLANEOUS: Binding of EGF to the receptor leads to
CC	dimerization, internalization of the EGF-receptor complex,
CC	induction of the tyrosine kinase activity, stimulation of cell DNA
CC	synthesis, and cell proliferation.
CC	-I- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC	or send an email to license@isb-sdb.ch).
CC	-----
DR	EMBL; X00588; CAA25240.1; -
DR	EMBL; U95089; AAB53063.1; -
DR	EMBL; U48722; AAC50802.1; -
DR	EMBL; U48723; AAC50804.1; -
DR	EMBL; U48724; AAC50796.1; -
DR	EMBL; U48725; AAC50797.1; -
DR	EMBL; U48726; AAC50798.1; -
OY	Query Match 19.3%; Score 721; DB 1; Length 1210;
OY	Best Local Similarity 41.3%; Pred. No. 2.4e-32;
OY	Matches 149; Conservative 51; Mismatches 129; Indels 30; Gaps 6;
DQ	36 ARGASTGVCTGDMDKRLRPLASPERHLMLRLHYGGCOVGNLELTLYPTNASLFOD 95
DQ	22 SRALEEKYCGSTSKNLQLGTGFDEHFSLDMFMFNCEVYLGNLEITTYVORNYDLSFLT 81
OY	96 IOEVGVVLIAHNVOYRVPLOLRIRVRGTOTLEDNYALVALVDNGDPUNNTTPVTGASBQG 155
DQ	82 IOEAVGVVLIALLNTERIPLENLQIIIRGMVMYENSVALVALSNVD-----ANKTG 131
OY	156 LRELQLRSLLTLKGCGLIQRNPOLCYDTJIMKDIFHKNNOLALTLDITNSRACHPCS 215
DQ	132 LKELEPMRNIOELIHGAVFSSNPALCNESIOWRDIVSSDFLSNMSPFOHNLGSCOKCD 191
OY	216 PMCKSGRCWGESSEDCOSLTRTVCAAGGA-RCKGPLPTDCCHEGQAACATGKHSHDCLAG 274
DQ	192 PSCPGSGWGAEENCKQLTKLICAKQGSGRKRGKSPSDCCHNCDDAACCTGPRESDCLVC 251

Qy	275	LHFNHSGICELHCPALVNTYNTDFSPMSNDEPRYFFGASVCYACVYNTLSTDPVSGAGGM	334
Db	232	RKFDEATCCKDTCCPLMLXNPTTYMDVYNPECKYSFEGATCYAKKCRNNVYVDHG-----	305
Qy	335	VHHRHRSSTSRGGDGLTGLPESEEARPSRLAPSEAGSDVFDGDGLGMCAKQLSL	393
Db	306	-----SCVRACGAD-----SYEMEDGDYRK-CKKCEGFCRKVCNG-IGIGFKDSLST	351
RESULT 6			
ERB4_HUMAN	ERB4_HUMAN	STANDARD:	PRT: 1308 AA.
AC	Q15303:		
DT	15-DEC-1998	(Rel. 37, Created)	
DT	15-DEC-1998	(Rel. 37, Last sequence update)	
DT	15-JUN-2002	(Rel. 41, Last annotation update)	
DE	Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112)		
DE	(p180erbB4) (Tyrosine kinase erbB-4 protein)		
GN	ERBB4 OR HER4.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID:9606;		
PN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM JM-A).		
RC	TISSUE-Breast carcinoma:		
RX	MEDLINE-93189574; PubMed-8383326;		
RA	Plozman G.D., Culouscou J.-W., Whitney G.S., Green J.M., Carlton G.W.,		
RA	Foy L., Neubauer M.G., Shoyab M.;		
RT	"Ligand-specific activation of HER4/p180erbB4, a fourth member of the		
RT	epidermal growth factor receptor family.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 90:1746-1750(1993).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORMS JM-A AND JM-B).		
RC	TISSUE-Fetal brain:		
RX	MEDLINE-97476287; PubMed-9334263;		
RA	Elenius K., Cortes G., Paul S., Choi C.J., Rlo C., Plozman G.D.,		
RA	Klagsbrun M.;		
RT	"A novel juxtamembrane domain isoform of HER4/erbB4. Isoform-specific		
RT	tissue distribution and differential processing in response to		
RT	phorbol ester.";		
RL	J. Biol. Chem. 272:26761-26768(1997).		
CC	-1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULINS, NRG-		
CC	2, NRG-3, HEPARIN-BINDING EGF-LIKE GROWTH FACTOR, BETACELLULIN AND		
CC	NRK. INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.		
CC	NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN.		
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein		
CC	tyrosine phosphate.		
CC	-1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE ERBB		
CC	RECEPTORS (POTENTIAL).		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.		
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: JM-A (SHOWN HERE) AND JM-B:		
CC	ARE PRODUCED BY ALTERNATIVE SPLICING. THE TWO ISOFORMS DIFFER		
CC	FUNCTIONALLY IN THEIR RESPONSE TO PHORBOL ESTER: JM-A IS PROCESSED		
CC	BUT NOT JM-B. SO, THEY RESPECTIVELY REPRESENT CLEAVABLE AND		
CC	NONCLEAVABLE FORMS OF THE RECEPTOR. BOTH ISOFORMS ARE EXPRESSED IN		
CC	CEREBELLUM, BUT ONLY THE JM-B ISOFORM IS EXPRESSED IN THE HEART.		
CC	-1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN BRAIN, HEART,		
CC	KIDNEY, IN ADDITION TO SKELETAL MUSCLE, PARATHYROID, CEREBELLUM,		
CC	PIUITARY, SPLEEN, TESTIS AND BREAST. LOWER LEVELS IN THYMUS,		
CC	LUNG, SALIVARY GLAND, AND PANCREAS.		
CC	-1- PM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE		
CC	RESIDUES.		
CC	-1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		

CC EMBL: L07868; AAB59446.1; -
 DR HSP: P11362; IEGK.
 DR Genew: HGNC:3432; ERBB4.
 DR MIM: 600543; -
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR Pfam: PF02757; YLP; 2.
 DR Prodom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00261; FU; 4.
 DR SMART: SM00219; TYRK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Alternative splicing.
 KW SIGNAL 1 25
 FT CHAIN 26 1308 POTENTIAL.
 FT DOMAIN 26 651 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-4.
 FT TRANSMEM 652 675 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 676 1308 POTENTIAL.
 FT DOMAIN 186 334 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 496 633 CYS-RICH.
 FT NP_BIND 718 985 PROTEIN KINASE.
 FT BINDING 724 732 ATP (BY SIMILARITY).
 FT ACN_SITE 843 843 BY SIMILARITY.
 FT DISULFID 189 197 BY SIMILARITY.
 FT DISULFID 193 205 BY SIMILARITY.
 FT DISULFID 213 221 BY SIMILARITY.
 FT DISULFID 217 229 BY SIMILARITY.
 FT DISULFID 230 238 BY SIMILARITY.
 FT DISULFID 234 246 BY SIMILARITY.
 FT DISULFID 249 258 BY SIMILARITY.
 FT DISULFID 262 289 BY SIMILARITY.
 FT DISULFID 293 304 BY SIMILARITY.
 FT DISULFID 308 323 BY SIMILARITY.
 FT DISULFID 326 330 BY SIMILARITY.
 FT DISULFID 503 512 BY SIMILARITY.
 FT DISULFID 507 520 BY SIMILARITY.
 FT DISULFID 523 532 BY SIMILARITY.
 FT DISULFID 536 552 BY SIMILARITY.
 FT DISULFID 555 569 BY SIMILARITY.
 FT DISULFID 559 577 BY SIMILARITY.
 FT DISULFID 580 589 BY SIMILARITY.
 FT DISULFID 593 614 BY SIMILARITY.
 FT DISULFID 617 625 BY SIMILARITY.
 FT DISULFID 633 633 BY SIMILARITY.
 FT MOD_RES 1162 1162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1188 1188 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1258 1258 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT MOD_RES 1284 1284 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 174 174 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 358 358 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 495 495 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 548 548 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 576 576 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VANSPLIC 626 648 NGPTSHDCTIYPMGHSFLPPIHA -> IGSSIEDICIGLMD (IN ISOFORM JM-B).
 SQ SEQUENCE 1308 AA; 146807 MW; 5E4AE80985D88761 CRC64;

Query Match 19.3%; Score 721; DB 1; Length 1308;
 Best Local Similarity 46.9%; Pred. No. 2,66-32;
 Matches 142; Conservative 42; Mismatches 107; Indels 12; Gaps 4;
 Db 22 LFFWIDRSVLAKELARGASQVCTGTDMLRLPASPTHLDMLHLYQGCQVYVGNLEL 81
 Db 7 LVMVWLLVAAGTV-QPDSQSVAGGTENKLSLSLSDLOQYRALRKYYENCEVVGNDLEI 65
 Db 82 TYLPINASLFLQDIOEVQGVYLLAHNOYROYPIQRLNIVGTQGLFEDENYALAVLDNDP 141
 Db 66 TSIEHNRDLSELSRVETGVYLVALNOFRLPLENELLIRKTYEDRYALALFLNLRK 125
 Db 142 LNTTPVVGASPGGLRELOLSLTFEILKGVLIORNPOLCQODTILMDIFHKNNQLALT 201
 Db 126 DGNF-----GLQELGLKNLLEILNGVYVDQNKFLCTADITHMODIYRNPMSULT 176
 Db 202 LIDTNRBRACHPCSPKSGKSGKMGESSEDCOSLPTTVGAGC-ARCKGPLEPTDCHEOCA 260
 Db 177 LVSTNGSSGCGRHKSCGTG-RQWGPTEHNCOTLRTVCAECODGRCYGPVYSDCCHRECA 235
 Db 261 AGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFESMPNPGRTTFGASCYTACP 320
 Db 236 GGCSPKPDIDCFACMNFDSGACVYQCPQTPVYNPPTFQLEHNFNAKYTGAFVCKCPH 295
 Db 321 NYL 323
 Db 296 NFV 298
 RESULT 7
 EGFR_MOUSE STANDARD; PRT: 1210 AA.
 AC 001279;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Epidermal growth factor receptor precursor (EC 2.7.1.112).
 GN EGFR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE-Liver;
 RX MEDLINE-93026370; PubMed-1408137;
 RA Avivi A., Skorecki K., Yajon A., Givol D.;
 RT "Promoter region of the murine fibroblast growth factor receptor 2
 (bek/KGFR) gene.";
 RL Oncogene 7:1957-1962(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; and CD-1; TISSUE-Liver, and Decidua;
 RX MEDLINE-93126380; PubMed-7678348;
 RA Parla B.C., Das S.K., Andrews G.K., Dey S.K.;
 RT "Expression of the epidermal growth factor receptor gene is regulated
 in mouse blastocysts during delayed implantation.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:55-59(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE-Liver;
 RX Hibbs M.L.;
 RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B6/C3; TISSUE-Liver;
 RX MEDLINE-94170986; PubMed-812525;
 RA Luetteke N.C., Phillips H.K., Qiu T.H., Copeland N.G., Earp H.S.,
 RA Jenkins N.A., Lee D.C.;
 RT "The mouse waved-2 phenotype results from a point mutation in the EGF
 receptor tyrosine kinase.";
 RL Genes Dev. 8:399-413(1994).

RN [5]
 RP SEQUENCE OF 1-714 FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-91232866; PubMed-2030916;
 RA Avila A., Lax I., Ullrich A., Schlessinger J., Givol D., Morse B.;
 RT "Comparison of EGF receptor sequences as a guide to study the ligand
 binding site.";
 RL Oncogene 6:673-676(1991).
 RN [6]
 RP SEQUENCE OF 969-1117 FROM N.A.
 RC STRAIN-CH;
 RA Elstner D.P., Serrero G.;
 RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE EGF RECEPTOR MEDIATES THE BIOLOGICAL SIGNAL OF EGF,
 AND ALSO OF TGF-ALPHA, AMPHIREGULIN, HEPARIN-BINDING EGF, GP30 AND
 VACCINIA VIRUS GROWTH FACTOR (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- MISCELLANEOUS: Binding of EGF to the receptor leads to
 dimerization, internalization of the EGF-receptor complex,
 induction of the tyrosine kinase activity, stimulation of cell DNA
 synthesis, and cell proliferation.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X78987; CAA55587.1; -
 DR EMBL: U03425; AAA17899.1; -
 DR EMBL: X59698; CAA42219.1; -
 DR EMBL: L06864; AAA53029.1; -
 DR EMBL: Z12608; CAA78249.1; -
 DR HSSP: P11362; IFGK.
 DR MGD: MGI:95294; Egfr.
 DR InterPro: IPR000494; EGFR_L_domain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF01030; Recep_L_domain; 2.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00261; FU; 3.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Receptor; Signal; Transferase;
 KM Tyrosine-protein kinase; ATP-binding; Phosphorylation; Repeat.
 FT CHAIN 1 24
 FT DOMAIN 25 1210
 FT TRANSMEM 647 670
 FT DOMAIN 671 1210
 FT REPEAT 75 300
 FT REPEAT 390 600
 FT DOMAIN 1028 1071
 FT DOMAIN 714 981
 FT NP_BIND 720 728
 FT BINDING 747 747
 FT ACT_SITE 839 839
 FT DISULFID 190 199
 FT DISULFID 194 207
 FT DISULFID 215 223
 FT DISULFID 219 231
 FT DISULFID 232 240
 FT DISULFID 236 248
 FT DISULFID 236 248

FT DISULFID 251 260 BY SIMILARITY.
 FT DISULFID 264 291 BY SIMILARITY.
 FT DISULFID 295 307 BY SIMILARITY.
 FT DISULFID 311 326 BY SIMILARITY.
 FT DISULFID 329 333 BY SIMILARITY.
 FT DISULFID 506 515 BY SIMILARITY.
 FT DISULFID 510 523 BY SIMILARITY.
 FT DISULFID 526 535 BY SIMILARITY.
 FT DISULFID 539 555 BY SIMILARITY.
 FT DISULFID 558 571 BY SIMILARITY.
 FT DISULFID 562 579 BY SIMILARITY.
 FT DISULFID 582 591 BY SIMILARITY.
 FT DISULFID 595 617 BY SIMILARITY.
 FT DISULFID 620 628 BY SIMILARITY.
 FT DISULFID 624 636 BY SIMILARITY.
 FT DISULFID 680 680 BY SIMILARITY.
 FT MOD_RES 1092 1092 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
 FT MOD_RES 1110 1110 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 1172 1172 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 1197 1197 PHOSPHORYLATION (BY SIMILARITY).
 FT CARBOHYD 128 128 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 528 528 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 19 19 C -> S (IN REF. 2).
 FT CONFLICT 539 539 C -> W (IN REF. 5).
 FT CONFLICT 991 991 L -> F (IN REF. 4).
 FT CONFLICT 1116 1117 HP -> DR (IN REF. 6).
 SQ SEQUENCE 1210 AA; 134853 MW; 690E20D46DF2D2F5 CRC64;
 Query Match 19.2%; Score 717.5; DB 1; Length 1210;
 Best Local Similarity 46.8%; Pred. No. 3; 7e-32;
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;
 DB 36 AAGAA--STOVCCTGTDKRLPASPETHLDMLVCGCOVONLETLVPTNASTFL 93
 DB 20 AAGAALEKRVCGTSGKRLTQTFEDHPLSLDQRMNCEVVLGNLEITYVQNRDLSFL 79
 QY 94 QDIQEVGVYVLAHNOVQVPLQRLRVCTQLEFEDYVALAVLDNGDPLNNTTPTVTCASP 153
 DB 80 KTIQEVAGVYVLAHNTVERIPLENLQIRGNALYENTYALALISN-----YGTNR 129
 QY 154 GGLRELQRLSTLEILNGGVLIQNRNPOLCYODTILMKDI-----FHKNNQALALTLIDNRSR 209
 DB 130 TGLRELPMRLQELILGAVRFSSNPNILCNDTQWRIDIVONVFSNMSMDL---OSHPS 185
 QY 210 ACHPSCPMGRCSCWSSSDCSLRTVTCAGCA--RCKGPLPTDCHEQCAAGTCPRK 268
 DB 186 SCRCPCSCNGSCWGGEGECCCKLRIICAOQCSHRRCRSRSDCHNCAGAGCTGPR 245
 QY 269 SDCLACLFHNSGICELHCPALVYNTDFESMPNPEGRTTGCASCVTACPVYLTVDG 328
 DB 246 SDCLVQKQFQDEATCKDTCPLMLYNTTYQMDVNEGKYSFGATGVKCRMYVYVDHG 305
 QY 329 S 329
 DB 306 S 306
 RESULT 8
 ID ERB4_RAT STANDARD; PRT; 1308 AA.
 AC Q62956; Q922N7;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Receptor protein-tyrosine kinase erbB-4 precursor (EC 2.7.1.112).
GN ERBB4 OR TYRO-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Heart;
RX MEDLINE=98221155; PubMed=9553078;
RA Zhao Y.-Y., Sawyer D.R., Baliga R.R., Opel D.J., Han X.,
RA Marchionni M.A., Kelly R.A.;
RT "Neuregulins promote survival and growth of cardiac myocytes.
RT Persistence of ErbB2 and ErbB4 expression in neonatal and adult
RT ventricular myocytes.";
RL J. Biol. Chem. 273:10261-10269(1998).
RN [2]
RP SEQUENCE OF 848-901 FROM N.A.
RC TISSUE-Sciatic nerve;
RX MEDLINE=9122560; PubMed=2025425;
RA Lai C., Lemke G.;
RT "An extended family of protein-tyrosine kinase genes differentially
RT expressed in the vertebrate nervous system.";
RL Neuron 6:691-704(1991).
RN [3]
RP SEQUENCE OF 1031-1198 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Spinal cord;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuregulins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration.";
RL J. Neurosci. 17:1642-1659(1997).
CC -1- FUNCTION: SPECIFICALLY BINDS AND IS ACTIVATED BY NEUREGULIN AND
CC NRG-2. HEPARIN-BINDING EGF-LIKE GROWTH FACTOR. HETEROLOGOUS AND
CC NOT ACTIVATED BY EGF, TGF-A, AND AMPHIREGULIN (BY SIMILARITY).
CC INTERACTION WITH THESE FACTORS INDUCES CELL DIFFERENTIATION.
CC CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HOMODIMER OR HETERODIMER WITH EACH OF THE OTHER ERBB
CC RECEPTORS (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE DEVELOPING
CC NERVOUS SYSTEM. EXHIBITS DISTINCT AND HIGHLY REGIONALIZED PATTERNS
CC OF EXPRESSION IN THE ADULT BRAIN, WHERE IT IS MAINLY FOUND IN THE
CC RETICULAR NUCLEUS OF THE THALAMUS. VERY LOW LEVELS IN KIDNEY, AND
CC HEART.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE
CC RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF041838; AAD08899.1; -
DR EMBL: U52531; AAC53051.1; -
DR HSSP: P11362; IFGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR004019; YLP_motif.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; Kinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR Pfam: PF02757; YLP; 2.
DR PRINTS: PD00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; FU; 4.

DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00101; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 25
FT CHAIN 26 1308
FT DOMAIN 26 651
FT TRANSMEM 652 675
FT DOMAIN 676 1308
FT DOMAIN 186 334
FT DOMAIN 496 633
FT DOMAIN 718 985
FT NP_BIND 724 732
FT BINDING 751 751
FT ACT_SITE 843 843
FT DISULFID 189 197
FT DISULFID 193 205
FT DISULFID 213 221
FT DISULFID 217 229
FT DISULFID 230 238
FT DISULFID 234 246
FT DISULFID 249 258
FT DISULFID 262 289
FT DISULFID 293 304
FT DISULFID 308 323
FT DISULFID 326 330
FT DISULFID 503 512
FT DISULFID 507 520
FT DISULFID 523 532
FT DISULFID 536 552
FT DISULFID 555 569
FT DISULFID 559 577
FT DISULFID 580 589
FT DISULFID 593 614
FT DISULFID 617 625
FT DISULFID 621 633
FT MOD_RES 1162 1162
FT MOD_RES 1188 1188
FT MOD_RES 1258 1258
FT MOD_RES 1284 1284
FT CARBOHYD 138 138
FT CARBOHYD 174 174
FT CARBOHYD 253 253
FT CARBOHYD 358 358
FT CARBOHYD 410 410
FT CARBOHYD 473 473
FT CARBOHYD 495 495
FT CARBOHYD 548 548
FT CARBOHYD 576 576
FT CARBOHYD 620 620
FT CONFLICT 1062 1062
FT CONFLICT 1080 1082
FT SEQUENCE 1308 AA; 146957 MW; D944BB0996A0B41 CRC64;
Query Match 19.28; Score 717; DB 1; Length 1308;
Best Local Similarity 47.28; Pred. No. 42e-33;
Matches 142; Conservative 41; Mismatches 106; Indels 12; Gaps 4;
QY 25 WLDKSVL-AKELRGAASSTOYCTGTDMKRLPASPEHLDMLRHLXGCGOYVGNLELY 83
DB 8 WVMGSLVAVARTVOPASQSVCACTENKLSLSLDEQGYBALRKYRVECEVVMGNETITS 67
QY 84 LPTNASLFDLDDIOEVQGYVLIANNQVYVQLQRLRYRGQLFEDNRYALVLDNGDPLN 143
DB 68 IEHRDLSFLRSIREYGVYVVALNQPFLPLNRIIRGKLYEDRYALALFVNYRKG 127
QY 144 NTFVTGASPGGLRELDRLSLTEILKGVLFQRPOLCYOPTITLMDKDFHNNQALATLI 203
DB 128 NF-----GLQDLGKNTLETILNGYVYDQKFLCYADTITMODIVIRPWSMNTLV 178
QY 204 DTNRSRACHPCSPMCKGSRGCESSDCCQSLTRTVACGCC-ARCKGPLPTDCCHEQCAAG 262

Query Match 18.6%; Score 697; DB 1; Length 1342;
 Best Local Similarity 44.6%; Pred. No. 5.4e-31;
 Matches 139; Conservative 45; Mismatches 98; Indels 30; Gaps 7;

16 LGFLEFLFWLDRSVLAKELARGAA---STOYCTGTDMKRLPASPETHLMDLRHLQGC 72
 10 LGLFLF-----SLARGSEVNSQAVCPCTLNGLSVTGDAENQOYLYKLYERC 56
 73 QVQGNLELTYLPTNNAISLFIQDQDEQGYVLIHANOYRQVPLQRLRYRGTOLFEDNYA 132
 57 EVVGNNEIYLTGHADISFLQWIREVGYLVAMNESTPLPLRLRYRGTQYDGF 116
 133 LAVDNGDPLNNTPTVTSAGSPGLRELOLRSLTEILKGVLIQRPOLCYODTIFMKDIF 192
 117 IFWV-----LKYNT-----NSSHALRQLRLTOLTELSGVYIEKNDKLCMDTIDMRDIY 167
 193 HKNQOLATLTDITNRSRACHPCSPMCKSGRCWGSSEDDQSLTRTVACGG-ARCKGFLP 251
 168 RDRD---AEIYVKDNGRCPCPCHEVCKG-RCWGPGESEDCQRLTITICAPQCNGHCFGNP 223
 252 TDCCHQCAACTGPKHSDCLACLHFNNSGICELHCPALVYNTDFESMNPERRTFG 311
 224 NQCHDECAAGCGSPQDTDFCAHRHFNDSGACVPRCPQPLVYNTLTLQLEBNPHTKYQYG 283
 312 ASCVTACPYNTL 323
 284 GCVVASCPHNFV 295

RESULT 10
 CSF2_HUMAN
 ID CSF2_HUMAN STANDARD; PRT; 144 AA.
 AC P04141;
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
 DE (colony-stimulating factor) (CSF) (Sargramostim) (Molgramostin).
 GN CSF2 OR GMCSF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_Taxid:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85242684; PubMed-3925454.
 RA Lee F., Yokota T., Otsuka T., Gemmell L., Larson N., Luh J.,
 RA Arai K.-I., Rennick D.;
 RT "Isolation of cDNA for a human granulocyte-macrophage
 RT colony-stimulating factor by functional expression in mammalian
 RT cells";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4360-4364(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86205844; PubMed-3486413;
 RA Kaushansky K., O'Hara P.J., Berkner K., Segal G.M., Hagen F.S.,
 RA Adamson J.W.;
 RT "Genomic cloning, characterization, and multilineage growth-promoting
 RT activity of human granulocyte-macrophage colony-stimulating factor";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3101-3105(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85298329; PubMed-3898082;
 RA Cantrell M.A., Anderson D., Cerretti D.P., Price V., McKereghan K.,
 RA Tuninski R.J., Mochizuki D.Y., Larsen A., Grabstein K., Cosman D.,
 RT "Cloning, sequence, and expression of a human granulocyte/macrophage
 RT colony-stimulating factor";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6250-6254(1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85218749; PubMed-3923623;
 RA Wong G.G., Wittek J.S., Temple P.A., Wilkens K.M., Leary A.C.,

RA Luxenberg D.P., Jones S.S., Brown E.L., Kay R.M., Orr E.C.,
 RA Shoemaker C., Golde D.W., Kaufman R.J., Hewick R.M., Wang E.A.,
 RA Clark S.C.;
 RT "Human GM-CSF: molecular cloning of the complementary DNA and
 RT purification of the natural and recombinant proteins";
 RL Science 228:810-815(1985).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-86030234; PubMed-3876930;
 RA Miyatake S., Otsuka T., Yokota T., Lee F., Arai K.-I.;
 RT "Structure of the chromosomal gene for granulocyte-macrophage colony
 RT stimulating factor: comparison of the mouse and human genes";
 RL EMBO J. 4:2561-2568(1985).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
 RA Davis C.A., Kader K., Miguel T., Pittluck S., Pollard M., Rojaski H.,
 RA Subramanian S., Martin C.H.;
 RT submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. AND VARIANTS ILE-115 AND THR-117.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RT submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE-92144609; PubMed-1737041.
 RA Kaushansky K., Lopez J.A., Brown C.B.;
 RT "Role of carbohydrate modification in the production and secretion of
 RT human granulocyte macrophage colony-stimulating factor in genetically
 RT engineered and normal mesenchymal cells";
 RL Biochemistry 31:1881-1886(1992).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE-92108420; PubMed-1837174.
 RA Dederichs K., Boone T., Karpus P.A.;
 RT "Novel fold and putative receptor binding site of
 RT granulocyte-macrophage colony-stimulating factor";
 RL Science 254:1779-1782(1991).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE-92235844; PubMed-1569568;
 RA Walter M.R., Cook W.J., Ealick S.E., Nagabhushan T.L., Trotta P.P.,
 RA Buga C.E.;
 RT "Three-dimensional structure of recombinant human granulocyte-
 RT macrophage colony-stimulating factor";
 RL J. Mol. Biol. 224:1073-1085(1992).
 CC [1]
 CC FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
 CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
 CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES.
 CC [2]
 CC SUBUNIT: MONOMER.
 CC [3]
 CC POLYMORPHISM: VARIANT ILE-117 MAY BE A RISK FACTOR FOR ATOPIC
 CC ASTHMA.
 CC [4]
 CC PHARMACEUTICAL: Available under the names Leukine (Immunex) and
 CC Leucomax (Novartis). Used in myeloid reconstitution following bone
 CC marrow transplant, bone marrow transplant engraftment failure or
 CC delay, mobilization and following transplantation of autologous
 CC peripheral blood progenitor cells, and following induction
 CC chemotherapy in older adults with acute myelogenous leukemia.
 CC [5]
 CC DATABASE: NME-Leukine; NOTE-Clinical information on Leukine;
 CC WWW="http://www.imunex.com/patient/pa02el.html".
 CC [6]
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC [7]
 CC EMBL: M13207; AAA96768.1; -;
 CC EMBL: M11734; AAA52122.1; -;
 CC EMBL: M11220; AAA52578.1; -;


```

DR EMBL: X03021; CAA26822.1; -
DR EMBL: M10653; AAA52121.1; -
DR EMBL: AC004511; AAC08707.1; -
DR EMBL: AF373868; AAK51563.1; -
DR PIR: A01853; FOHUGM.
DR PIR: C24636; C24636.
DR PIR: A25169; A25169.
DR PDB: 1CSG; 31-JAN-94.
DR PDB: 2GMF; 08-NOV-96.
DR Genew: HGNC:2434; CSF2.
DR MIM: 138960; -.
DR InterPro: IPR000773; GM_CSF.
DR Pfam: PF01109; GM_CSF.1.
DR PRINTS: PR00693; GMCSFAC.
DR PRODOM: PD007349; GM_CSF.1.
DR SMART: SM00040; CSF2.1.
DR PROSITE: PS00702; GM_CSF.1.
DR KEGG: Glycokine; Growth factor; Glycoprotein; Signal; 3D-structure;
KM Polymorphism; Pharmaceutical.
FT SIGNAL 1 17
FT CHAIN 18 144
FT FT
FT DISULFID 71 113
FT FT
FT FT 105 138
FT CARBOHYD 22 22
FT CARBOHYD 24 24
FT CARBOHYD 26 26
FT CARBOHYD 27 27
FT CARBOHYD 44 44
FT CARBOHYD 54 54
FT VARIANT 115 115
FT FT
FT VARIANT 117 117
FT FT
FT TURN 25 27
FT HELIX 30 44
FT TURN 45 45
FT HELIX 50 54
FT STRAND 56 60
FT HELIX 72 81
FT TURN 82 82
FT HELIX 85 103
FT TURN 104 104
FT STRAND 115 119
FT HELIX 120 131
FT TURN 132 133
SQ SEQUENCE 144 AA; 16295 MW; 75DIE50506BCA7A8 CRC64;

Query Match 18.0%; Score 674; DB 1; Length 144;
Best Local Similarity 99.2%; Pred. No. 9.5e-31;
Matches 127; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 548 AAFARSPSTQPMENHNAIOEARRLNLSRDTAAENETVEYISEMFDLOEPTCLOTRL 607
DB 17 SARARSPSTQPMENHNAIOEARRLNLSRDTAAENETVEYISEMFDLOEPTCLOTRL 76
OY 608 ELYKQGLRSLTKLKGPLTMASHYKQPCPPETSCATQIITFESEKEMDKDFLVIPF 667
DB 77 ELYKQGLRSLTKLKGPLTMASHYKQPCPPETSCATQIITFESEKEMDKDFLVIPF 136
OY 668 DCWEPEQOE 675
DB 137 DCWEPEQOE 144

RESULT 11
ERR3_RAT
ID ERR3_RAT STANDARD: PRT; 1339 AA.
AC Q62799; Q62955;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-3 precursor (EC 2.7.1.112)

```

```

DE (c-erbB3).
GN ERB3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Liver;
RX MEDLINE=86096535; PubMed=8522190;
RA Heller N.J., Kim H.-H., Greaves C.H., Sterke S.L., Koland J.G.;
RT "Cloning of the rat ErbB3 cDNA and characterization of the
RT recombinant protein.";
RL Gene 165:279-284(1995).
RN [2]
RP REVISIONS TO 85: 513 AND 565.
RA Heller N.J., Koland J.G.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 922-1097 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-sclatic nerve;
RX MEDLINE=97184212; PubMed=9030624;
RA Carroll S.L., Miller M.L., Frohner P.W., Kim S.S., Corbett J.A.;
RT "Expression of neuroligins and their putative receptors, ErbB2 and
RT ErbB3, is induced during Wallerian degeneration.";
J. Neurosci. 17:1642-1659(1997).
CC -1- FUNCTION: BINDS AND IS ACTIVATED BY NEUREGULINS AND NTAK.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBUNIT: HETERODIMER WITH EACH OF THE OTHER ERBB RECEPTORS
CC (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DOMAIN: THE CYTOPLASMIC PART OF THE RECEPTOR MAY INTERACT WITH THE
CC SH2 OR SH3 DOMAINS OF MANY SIGNAL-TRANSDUCING PROTEINS.
CC -1- PTM: LIGAND-BINDING INCREASES PHOSPHORYLATION ON TYROSINE RESIDUES
CC AND PROMOTES ITS ASSOCIATION WITH THE P85 SUBUNIT OF
CC PHOSPHATIDYLINOSITOL 3-KINASE.
CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U29339; AAC28498.2; -
DR EMBL: U52530; AAC53050.1; -
DR HSSP: P11362; IFGK.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR Pfam: PF00757; Furin-like.1.
DR Pfam: PF01030; Recep_L_domain.2.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase.1.
DR SMART: SM00261; FU_5.
DR SMART: SM00219; TYRKC.1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
DR Transmembrane: Glycoprotein; Multigene family; Receptor; Signal;
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT SIGNAL 1 19
FT CHAIN 20 1339
FT DOMAIN 20 643
FT TRANSMEM 644 662
FT DOMAIN 663 1339
FT DOMAIN 183 259
FT DOMAIN 707 964
PROTEIN KINASE.

```



```

FT DISULFID 300 311 BY SIMILARITY.
FT DISULFID 315 330 BY SIMILARITY.
FT DISULFID 333 337 BY SIMILARITY.
FT DISULFID 504 513 BY SIMILARITY.
FT DISULFID 508 521 BY SIMILARITY.
FT DISULFID 524 533 BY SIMILARITY.
FT DISULFID 537 533 BY SIMILARITY.
FT DISULFID 556 569 BY SIMILARITY.
FT DISULFID 560 577 BY SIMILARITY.
FT DISULFID 593 615 BY SIMILARITY.
FT DISULFID 618 626 BY SIMILARITY.
FT DISULFID 622 634 BY SIMILARITY.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1167 AA, 129934 MM, 4793E0749DC1D55A CRC64.

```

```

Query Match 17.5%; Score 655.5; DB 1; Length 1167;
Best Local Similarity 41.2%; Pred. No. 8.6e-29;
Matches 142; Conservative 46; Mismatches 130; Indels 27; Gaps 9;

```

```

OY 43 OVCSTGDMKRLRSPETHLDMLRHLVGGVGVGNELTYLPNASTSLFDIOIOVOGY 102
DB 30 KVCOSTGNOMTM---LDNHILKMKKMSGCVNVLNLEITTYOENOLSLFQSIQDEGXY 86
OY 103 VLIANOVROVPIORLRIIVGTOLFEEDNYALAVLNDGDPINNTTPYVAGSGRLRELQRL 162
DB 87 VLIAMNEVSTIPLVNLRLRIGONLYEGNFTLVMSNYOK--NPSSP--DYVQVGLKQQLS 143
OY 163 SLPEILKGVLIORNPOLCVDOTILMKDIFRKNQOLALTLIDTRKSRACPCSMCGSR 222
DB 144 NLEELISGVKAVSHNPLLCVNETIMWDIVDKTSNPTMNLIPHAEEQCKCKCHGVNGS 203
OY 223 CMGESSEDCOSLRTVYVAGGC--ARCKGPLPTDCHEOCAGACTGPKHSDCLACLFHNSG 281
DB 204 CMAPGPEHCCKFTLLCAEQCNRRCKPRKPIDCNEHCAGGCTGPRATDCLACDFDNDG 263
OY 282 ICLHPCALVTYNTDFESHPNPEGRYTFGASCYVACPYNLSTDVSGAGGVVHHRHS 341
DB 264 TCKDTCPKPIYDISHQVVDNPNIKYTFGACVACVKECPSNVYVTE-----GACV-----RS 314
OY 342 SSTRSGGDLTLGLPSEEAERPSRLAPSGAGSDVFDGDLGKMA 386
DB 315 CSA-----GMLEVDENGKRS--CKPCDGVCPKVCDDG--IGIGS 348

```

```

RESULT 13
CSF2_SHEEP
ID CSF2_SHEEP STANDARD; PRT; 144 AA.
AC P28773;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).
GN CSF2.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92039044; PubMed=1937025;
RA McInnes C.J., Haig M.C.K.;
RT "Cloning and expression of a cDNA encoding ovine

```

```

RT granulocyte-macrophage colony-stimulating factor.";
RL Gene 105:275-279(1991).
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ish-sib.ch/announce/
CC or send an email to license@ish-sib.ch).
CC -----
CC EMBL: X53561; CA37632.1; -
CC PTR; JH0469; JH0469.
CC DR HSSP; P04141; ZGMF.
CC DR InterPro; IPR000773; GM_CSF.
CC DR Pfam; PF01109; GM_CSF.1.
CC DR PRINTS; PR00693; GMCSFACOR.
CC DR PRODOM; PD007349; GM_CSF.1.
CC DR SMART; SM00040; CSF2; 1.
CC DR PROSITE; PS00702; GM_CSF.1.
CC KW Cytokine; Growth factor; Glycoprotein; Signal.
CC FT SIGNAL 1 17 BY SIMILARITY.
CC FT CHAIN 18 144 GRANULOCYTE-MACROPHAGE COLONY-STIMULATING
CC FT FACTOR.
CC FT DISULFID 71 113 BY SIMILARITY.
CC FT DISULFID 103 138 BY SIMILARITY.
CC FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 144 AA, 16318 MM, ABAAC873B580008 CRC64;

```

```

Query Match 14.7%; Score 549; DB 1; Length 144;
Best Local Similarity 79.7%; Pred. No. 6.4e-24;
Matches 102; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

```

```

OY 548 AAPARSPSPTOPMEVNAIOEARRLNLSRDPAENKEVEVISENFDLQEPICLQTRL 607
DB 17 SAPTRGSSPYTRPQHVDAIKELSLNDSTDFAAVDEVEVYSEMFDSQEPICLQTRL 76
OY 608 ELVYQGRGSLTKKGLPTLMASHYKCHCPPTSCATQITPESKENVKDLVYIPF 667
DB 77 ELVKQGRGSLTSLTGLTWMASHYKRCPTQETSCETQITPESKENVKDLVYIPF 136
OY 668 DCMPEVOE 675
DB 137 DCMPEVOK 144

```

```

RESULT 14
EGFR_DROME
ID EGFR_DROME STANDARD; PRT; 1426 AA.
AC P04412; O61601; Q9Y2G0; P81668;
DT 13-AUG-1987 (Rel. 05, Created)
DT 15-AUG-1987 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermal growth factor receptor precursor (EG 2.7.1.112) (Egfr)
DE (Gurken receptor) (Tropo protein) (Drosophila relative of ERBB).
GN EGFR OR TOP OR C-ERBB OR DER OR CG10079.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscophora; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS TYPES I AND II).
RX MEDLINE=94350209; PubMed=8070664;
RA Clifford R., Schubach T.;
RT "Molecular analysis of the Drosophila EGF receptor homolog reveals
RT that several genetically defined classes of alleles cluster in
RT subdomains of the receptor protein.";

```

RL Genetics 137:553-550(1994).
 RN [2]
 RP REVISIONS.
 RA Clifford R., Schupbach T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE-85134611; PubMed-2982499;
 RA Llywne E., Glazer L., Segal D., Schlesinger J., Shilo B.-Z.;
 RT "The Drosophila EGF receptor gene homolog: conservation of both
 RT hormone binding and kinase domains.";
 RL Cell 40:599-607(1985).
 RN [4]
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ALTERNATIVE SPLICING.
 RC STRAIN-Oregon-R; TISSUE-Embryo;
 RX MEDLINE-87002474; PubMed-3093080;
 RA Schejter E.D., Segal D., Glazer L., Shilo B.-Z.;
 RT "Alternative 5' exons and tissue-specific expression of the
 RT Drosophila EGF receptor homolog transcripts.";
 RL Cell 46:1091-1101(1986).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I), TISSUE SPECIFICITY, AND MUTATION
 RP ANALYSIS.
 RX MEDLINE-99102120; PubMed-9882502;
 RA Leachman A.M., Yu S.-Y., Katz J., Baker N.E.;
 RT "Several levels of EGF receptor signaling during photoreceptor
 RT specification in wild-type, Ellipse, and null mutant Drosophila.";
 RL Dev. Biol. 205:129-144(1999).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM TYPE I).
 RC STRAIN-Berkley;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Landell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beson K.Y., Bencos P.V., Bernier B.P., Bhandari D., Boltskov S.,
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Pollara J., Puri V., Reese M.G.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollara J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schejter E., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [7]
 RP SEQUENCE OF 959-1078 FROM N.A.
 RC STRAIN-Daekwanryeong;

RX MEDLINE-85137938; PubMed-2983232;
 RA Wadsworth S.C., Vincent W.S. III, Bilodeau-Wentworth D.;
 RT "A Drosophila genomic sequence with homology to human epidermal
 RT growth factor receptor.";
 RL Nature 314:178-180(1985).
 RN [8]
 RP SEQUENCE OF 1133-1137; 1155-1167 AND 1209-1216 FROM N.A., AND MUTATION
 RP ANALYSIS.
 RX MEDLINE-92038942; PubMed-1936959;
 RA Raz E., Schejter E.D., Shilo B.Z.;
 RT "Interallelic complementation among DER/Elb alleles: implications for
 RT the mechanism of signal transduction by receptor-tyrosine kinases.";
 RL Genetics 129:191-201(1991).
 RN [9]
 RP REVIEW.
 RX MEDLINE-97248481; PubMed-9094709;
 RA Perrimon N., Perkins L.A.;
 RT "There must be 50 ways to rule the signal: the case of the Drosophila
 RT EGF receptor.";
 RL Cell 89:13-16(1997).
 CC -1- FUNCTION: BINDS TO FOUR LIGANDS: SPITZ, GURKEN, VEIN AND ARGOS,
 CC WHICH IS AN ANTAGONIST. TRANSDUCES THE SIGNAL THROUGH THE RAS-RAP-
 CC MAPK PATHWAY. INVOLVED IN A MYRIAD OF DEVELOPMENTAL DECISIONS.
 CC CRITICAL FOR THE PROLIFERATION OF IMAGINAL TISSUES, AND FOR THE
 CC DETERMINATION OF BOTH THE ANTERO-POSTERIOR AND DORSO-VENTRAL
 CC POLARITIES OF THE OOCYTE. IN THE EMBRYO, PLAYS A ROLE IN THE
 CC ESTABLISHMENT OF VENTRAL CELL FATES, MAINTENANCE OF AMNIOSEROSA
 CC AND VENTRAL NEUROECTODERMAL CELLS, GERM BAND REPACTION, CELL FATE
 CC SPECIFICATION IN THE CENTRAL NERVOUS SYSTEM AND PRODUCTION OF
 CC CUTICLE.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: ISOFORMS TYPE I AND TYPE II ARE TYPE I
 CC MEMBRANE PROTEINS. ISOFORM TYPE III IS A TYPE III MEMBRANE
 CC PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: TYPE I (SHOWN HERE), TYPE II AND
 CC TYPE III. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBICITOUSLY EXPRESSED IN EMBRYOS, IN LARVAE,
 CC UNIFORM EXPRESSION IS SEEN IN WING DISKS, GENITAL DISK, ANLAGEN OF
 CC TESTIS AND OVARY, AND BRAIN CORTEX. IN EYE-ANTENNA DISK, HIGHEST
 CC EXPRESSION IS ANTERIOR TO MORPHOGENETIC FURROW, LEVELS REMAIN HIGH
 CC IN PHOTORECEPTOR PRECURSOR CELLS. THIS PATTERN IS REVERSED IN
 CC POSTERIOR EYE DISK. IN ADULTS EXPRESSION IS HIGH IN BRAIN CORTEX
 CC AND THORACIC AND ABDOMINAL GANGLIA.
 CC -1- SIMILARITY: BELONGS TO THE EGF RECEPTOR FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF052754; AAC08536.1; -
 CC EMBL; AF052753; AAC08536.1; JOINED.
 CC EMBL; AF052752; AAC08535.1; -
 CC EMBL; AF052752; AAC08535.1; JOINED.
 CC EMBL; K03054; AAA51462.1; -
 CC EMBL; K03417; AAA51460.1; -
 CC EMBL; K03416; AAA50965.1; -
 CC EMBL; K03418; AAA51461.1; -
 CC EMBL; AF109077; AAD26134.1; -
 CC EMBL; AF109078; AAD26132.1; -
 CC EMBL; AF109082; AAD26132.1; JOINED.
 CC EMBL; AF109078; AAD26133.1; -
 CC EMBL; AF109084; AAD26133.1; JOINED.
 CC EMBL; AF109079; AAD26130.1; -
 CC EMBL; AF109081; AAD26130.1; JOINED.
 CC EMBL; AF109079; AAD26131.1; -
 CC EMBL; AF109083; AAD26131.1; JOINED.
 CC EMBL; AF109080; AAD26135.1; -
 CC EMBL; AF109080; AAD26135.1; -
 CC EMBL; AF003454; AAF46732.1; -

```

DR EMBL: X02293; CAA26157.1; -
DR EMBL: X78920; CAA55523.1; -
DR EMBL: X78918; CAA55521.1; -
DR EMBL: X78919; CAA55522.1; -
DR PIR: A00640; GQFE.
DR HSP: P11362; 1FGK.
DR FlyBase: FBgn0003731; Egfr.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase.1.
DR Pfam: PF00757; Furin-like.1.
DR Pfam: PF01030; Recep_L_domain.2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase.1.
DR SMART: SM00261; Fu; 7.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP.1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR.1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM.1.
DR Transmembrane; Glycoprotein; Receptor; Phosphorylation; Transferase;
KW Tyrosine-protein kinase; ATP-binding; Signal; Alternative splicing;
KM developmental protein.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 1426 EPIDERMAL GROWTH FACTOR RECEPTOR.
FT DOMAIN 31 868 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 869 889 POTENTIAL.
FT DOMAIN 890 1426 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 938 1198 CYTOPLASMIC (POTENTIAL).
FT NE_BIND 944 952 ATP (BY SIMILARITY).
FT BINDING 971 971 ATP (BY SIMILARITY).
FT ACT_SITE 1063 1063 BY SIMILARITY.
FT MOD_RES 902 902 PHOSPHORYLATION (BY PKC) (BY SIMILARITY).

Query Match 14.3%; Score 536; DB 1; Length 1426;
Best Local Similarity 36.8%; Pred. No. 3.6e-22;
Matches 112; Conservative 44; Mismatches 118; Indels 30; Gaps 7;

OY 34 ELARGASTVCTGTDMKRLPASPETHLMDRLHYGCGVGNGLTYLPT-NASLSF 92
DB 95 EFVKG---KICIGTKSRLSVPSNKEHYNRLRDYNTGYNQNLTLMPNENLDLSF 150
OY 93 LODIOEVGVLANHNOVROPRLRIIVGTOLF-----EDNVALVLDNGPLNNTTP 147
DB 151 LDNIREVTGGLISHDVKKVPEPKLIQINGRLFLSLVEEKVALTV----- 198
OY 148 VTGASPGGLRELRLRSLEILKGGVLTQRPOLCYODTILMKDIFHKNQDLATLIDTNR 207
DB 199 -----TYSKMTLELPDLADVLNGOVGFHNNYNLCHMTIOMSELVSGTDAIYNYDTAP 254
OY 208 SRACHPCSPKCGSRGSESSDOQLTRTVCAAGCA--RCKGPLPTDCHECACAGCTG 265
DB 255 ERECPKHCSECTHG-CWGESEPKNQKFSKLTSPQAGCGYCGPKPRECHLCAGCTG 313
OY 266 PKHSDCLACLPFHNSGICELHCPALTYNNDTESMNPREGRTTFCAGCTACPYNTLST 325
DB 314 PTQNDCTIACNPFDEAVSKCECPMRKRYNPTTYVLETNPBGKYAVGATCVKECP-GHLRL 372
OY 326 DVGS 329
DB 373 DNCA 376

RESULT 15
CSF2_CEREL STANDARD: PRT: 144 AA.
AC P51748:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Granulocyte-macrophage colony-stimulating factor precursor (GM-CSF)
DE (Colony-stimulating factor) (CSF).

```

```

GN CSF2.
OS Cervus elaphus (Red deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9660;
RN
RP SEQUENCE FROM N.A.
RA Lockhart E.A.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOKINE THAT STIMULATES THE GROWTH AND DIFFERENTIATION
CC OF HEMATOPOIETIC PRECURSOR CELLS FROM VARIOUS LINEAGES, INCLUDING
CC GRANULOCYTES, MACROPHAGES, EOSINOPHILS AND ERYTHROCYTES (BY
CC SIMILARITY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U14392; AAA21439.1; -
DR HSP: P04141; 2GMF.
DR InterPro: IPR000773; GM_CSF.
DR Pfam: PF01109; GM_CSF.1.
DR PRINTS: PR00693; GMCSFACOR.
DR ProDom: PD007349; GM_CSF.1.
DR SMART: SM00040; CSF2.1.
DR PROSITE: PS00702; GM_CSF.1.
KW Cytokine; Growth factor; Glycoprotein; Signal.
FT SIGNAL 1 17 BY SIMILARITY.
FT CHAIN 18 144 GRANULOCYTE-MACROPHAGE COLONY-STIMULATING
FT FT FACTOR.
FT DISULFID 71 113 BY SIMILARITY.
FT DISULFID 105 138 BY SIMILARITY.
FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 144 AA; 16283 MW; 1F5FF5FD03C94394 CRC64;

Query Match 14.1%; Score 529; DB 1; Length 144;
Best Local Similarity 76.6%; Pred. No. 7.9e-23;
Matches 98; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 548 AAPRSPSPSTQPEWHYNAIOEARRLNLSRPAENKNETVEVISEMFDQEPCTQTRL 607
DB 17 SAPRSPSPPTTRPQOHDAIKELSLNHSSDRAVNNETVEVISEMFDQEPCTQTRL 76
OY 608 ELYKQGLRGLTKLKGELTMASHYKQCPPTPETSCATGIIIFESKENVLKDPLVIFP 667
DB 77 KLYKQGLRGLSLTSLGSLTGMARHYEQHCPPTQETSCETGIIIFKSKENVLKDPLIFP 136
OY 668 DCWEPPGE 675
DB 137 DCWEPPAK 144

Search completed: April 28, 2003, 13:40:04
Job time : 21.4231 secs

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:38:09 ; Search time 25.3846 Seconds

(without alignments)
2613.108 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739
Sequence: 1 MRAPPLLAASLSGLF.....EPVQEGAPRRRAAHNNHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

PIR_73:.*
1: p1r1:.*
2: p1r2:.*
3: p1r3:.*
4: p1r4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2405.5	64.3	1255	1 A24571	protein-tyrosine k
2	1950.5	52.2	1260	1 TVRTNU	protein-tyrosine k
3	1935	51.8	1254	2 I48161	p-185 precursor
4	755.5	20.2	1223	1 TVCHLV	epidermal growth f
5	754.5	20.2	527	2 A42032	epidermal growth f
6	721	19.3	1210	1 GQHUE	epidermal growth f
7	721	19.3	1308	2 A47253	epidermal growth f
8	717.5	19.2	1210	2 A53183	epidermal growth f
9	714.5	19.1	644	2 A56325	epidermal growth f
10	697	18.6	1342	2 A36223	kinase-related tra
11	674	18.0	144	1 FQHUCM	granulocyte-macrop
12	662	17.7	1339	2 JC4387	epidermal growth f
13	655.5	17.5	1166	1 S06142	protein-tyrosine k
14	549	14.7	144	2 JH0469	granulocyte-macrop
15	545	14.6	144	1 A61632	granulocyte-macrop
16	536	14.3	843	2 A27131	epidermal growth f
17	481.5	12.9	143	1 FQBOGM	granulocyte-macrop
18	476.5	12.7	144	2 A49366	granulocyte-macrop
19	441	11.8	127	2 I46269	granulocyte-macrop
20	404.5	10.8	1323	2 E88257	protein let-23 (lm
21	404.5	10.8	1374	2 S70712	protein-tyrosine k
22	390.5	10.4	1369	2 S70713	protein-tyrosine k
23	375.5	10.0	1330	1 GQFPE	epidermal growth f
24	374	10.0	153	1 FQMSGM	granulocyte-macrop
25	341	9.1	366	2 D45558	epidermal growth f
26	341	9.1	1717	2 A45558	epidermal growth f
27	331	8.9	333	2 B45558	epidermal growth f
28	331	8.9	342	2 C45558	epidermal growth f
29	317.5	8.5	1363	2 T43220	insulin-like growt

30	273.5	7.3	1477	2 T18534	protein-tyrosine k
31	270.5	7.2	2101	2 S57245	insulin receptor (
32	270.5	7.2	2148	1 A56081	insulin receptor
33	256.5	6.9	1382	1 INHUR	insulin receptor p
34	252.5	6.8	1372	2 A34157	insulin receptor p
35	252	6.7	1607	2 T43212	insulin-like growt
36	249	6.7	1383	2 A36080	insulin receptor p
37	246.5	6.6	1300	2 A36502	insulin receptor p
38	246	6.6	1367	1 IGHUR1	insulin-like growt
39	238	6.4	340	2 B47417	insulin receptor-t
40	235.5	6.3	183	2 JH0803	tyrosine kinase re
41	235.5	6.3	1371	2 A33837	insulin-like growt
42	231.5	6.2	1268	2 B36502	insulin receptor-t
43	223.5	6.0	329	2 A48805	insulin-like growt
44	207	5.5	1390	2 T30346	insulin receptor h
45	188	5.0	1846	2 T42047	insulin receptor h

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: C-erb-B-2 protein precursor; kinase-related transforming protein e
C:Species: Homo sapiens (man)
C>Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999
C:Accession: A24571; A23491; A44188; B44188; I55509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Salto, T
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human C-erb-B-2 gene to epidermal growl
A:Reference number: A24571; MUID:66118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1235 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, C-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A23491; MUID:86016729; PMID:2955967
A:Accession: A23491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M1787; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Conssens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALI', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M17730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 837-909 <REX>
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAI>

A:Cross-references: GB:M16792; NID:9183983; PIDN:AAA58637.1; PID:9553332
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERBB2; NGL; NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-17q21.1
 A:Introns: 25/1; 75/3; 147/1; 883/3
 A>Note: the list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-125/Domain: signal sequence #status predicted <SIG>
 F:22-125/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
 F:22-653/Domain: extracellular #status predicted <EXT>
 F:70-304/Domain: EGF receptor extracellular domain repeat <E1>
 F:395-605/Domain: EGF receptor extracellular domain repeat <E2>
 F:654-675/Domain: transmembrane #status predicted <INT>
 F:676-125/Domain: intracellular #status predicted <INT>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 F:68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:753/Active site: lys #status predicted
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 64.3%; Score 2405.5; DB 1; Length 1255;
 Best Local Similarity 41.1%; Pred. No. 9, 7e-126;

Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGASTQVCTGTMKLRPLASPEETHLDMRLHYOGCQVVOGNETLYLPTNALSFLQ 94
 DB 16 LPPGASTQVCTGTMKLRPLASPEETHLDMRLHYOGCQVVOGNETLYLPTNALSFLQ 75
 QY 95 DIOVGQVYLIAHNOVQVLOLRIRYRGOLPEDNALVLDNGDPLNNTPTTGASRG 154
 DB 76 DIOVGQVYLIAHNOVQVLOLRIRYRGOLPEDNALVLDNGDPLNNTPTTGASRG 135
 QY 155 GLRELQRLSLTEILKGGVLLQIRNPQLCYODTILMKDIFRHNQALTLIDITNRSRACHP 214
 DB 136 GLRELQRLSLTEILKGGVLLQIRNPQLCYODTILMKDIFRHNQALTLIDITNRSRACHP 135
 QY 215 SPCKGRCRGESSEDCQSLTRTYCAGGACRCARCPRLTDCHECAGACGPKRSDCLAC 274
 DB 196 SPCKGRCRGESSEDCQSLTRTYCAGGACRCARCPRLTDCHECAGACGPKRSDCLAC 255
 QY 275 LHFHSGICELHCPALVTYNTDFESMPNEGRRTFGASCVTACPYLYLSTDVGSCTLVC 329
 DB 256 LHFHSGICELHCPALVTYNTDFESMPNEGRRTFGASCVTACPYLYLSTDVGSCTLVC 315
 QY 330 ----- 329
 DB 316 PLNHQEVTAADGTORCEKSCPCARVCYGLMEHLREVRVAVTSANTIOFAGCKKIFGSLA 375
 QY 330 ----- 329
 DB 376 FLPESEFDGDPASNTAPLQPOLOVFTLEETIGLYLISAMPDSLPLDSVFNQIVIRGRI 435
 QY 330 ----- 329
 DB 436 LHHGANSVLTLQGLGISWGLRSLRELGLALIHNTHTLFCFHTVPWDQLFRNPQALLH 495
 QY 330 ----- 329
 DB 496 TANRPDECEVGEGLACHQCLARGHCWGPRTQCVNCSQFLRGOECVEECRVLDGLPREYV 555
 QY 330 ----- 329
 DB 556 NARHCLPCHPECOPIONGSVTCFGEPAQCVAACHYKDPFCVACRCPGVKPLSLYPMWK 615
 QY 330 ----- 329

DB 616 FPBEAGACQCPINCTHSCVDLDDKCCPAQRASPLTSTISAVVGLLVVLCVFGIL 675
 QY 330 ----- 329
 DB 676 KRQOKIRKTYMRLQETELVEPLTPSGAMPQAOIRILKETELRKVKVLSGAGFTVY 735
 QY 330 ----- 329
 DB 736 KGIWIPDENVKIPVAIKVLRNTSPKANKELDEAVVMAGVSPYVRLGLCTSTVQ 795
 QY 330 ----- 329
 DB 796 LVTLQPLPYGCLDHYENRGRSLQSOLLNMCQIANGMSYLEDRVLRDIAARNVLYKS 855
 QY 330 ----- 329
 DB 856 PNHKIDPGLARLIDIDETEHADGKVPDKMALESILRRFTHQSDVKYGVYVWEL 915
 QY 330 ----- 329
 DB 916 MTEGAKPYDGIPIAREIPLDLEKEGRLLPQPICTIDVYIMKCMIDSECRPRRELVS 975
 QY 330 ----- 329
 DB 976 FSRMARDPQRFVVIQNDLGPASPLDSTFYRSLLEDMDKDLVDAEYLVPOQGFCCPD 1035
 QY 330 ----- 329
 DB 1036 APGAGGVHRRHSSSTRSGGDLTLGLEPSEEARSPPLAPSEGASDVFDGDLGMA 1095
 QY 388 KGLQSLPTHDPSLQRYSEDPTVPLPSETDGYAAPLTCSPQPEYVNOVDVPPSPREG 447
 DB 1096 KGLQSLPTHDPSLQRYSEDPTVPLPSETDGYAAPLTCSPQPEYVNOVDVPPSPREG 1155
 QY 448 PLPAARAGATLERRAKTLRSGKGVYKDVAFGAVENPEYLPPOGGAAPPPAPFSP 507
 DB 1156 PLPAARAGATLERRAKTLRSGKGVYKDVAFGAVENPEYLPPOGGAAPPPAPFSP 1215
 QY 508 AFDNLVYWDQDPPERGAPSTFKGTPTAENPEYLGDPV 546
 DB 1216 AFDNLVYWDQDPPERGAPSTFKGTPTAENPEYLGDPV 1254

RESULT 2

TVRTNU

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Dec-1988 #sequence-revision 31-Dec-1988 #text-change 11-Jun-1999

C:Accession: A24562; A61204

R:Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A:title: The neu oncogene encodes an epidermal growth factor receptor-related protein

A:Reference number: A24562; MUID:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 <BAR>

A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746

R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohe

Artile: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals n

2-thiazolyl]formamide or N-methyl-N-nitrosourea.

A:Reference number: A61204; MUID:92035293; PMID:1682063

A:Accession: A61204

A:Molecule type: DNA

A:Residues: 637-663 'V', 665-702 <MAS>

A>Note: authors translated the codon GCA for residue 25 as Val

C:Genetics:

A:Gene: neu

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

F:658-680/Domain: transmembrane #status predicted <TMN>
 F:723-988/Domain: protein kinase homology <KIN>
 F:721-739/Region: protein kinase ATP-binding motif
 F:71,191,263,535,576,634/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 52.28; Score 1950.5; DB 1; Length 1260;
 Best Local Similarity 34.18; Pred. No. 1.6e-100;
 Matches 433; Conservative 24; Mismatches 72; Indels 741; Gaps 3;

QY 6 LLLAASLSLGLFLLFWLDRSLAKELAGASTOVCTGTDMLRLPASPETHIDML 65
 Db 2 IIMELAMCRMGFLAL-----LPRGIAGTQVCGTDMKRLRASPETHIDML 49
 QY 66 RHLYOGGVVGNLELTPYLPNALSFLQDIQVGVYLAHNOVROYDRLRIVNGTO 125
 Db 50 RHLYOGGVVGNLELTPYLPNALSFLQDIQVGVYLAHNOVROYDRLRIVNGTO 109
 QY 126 LFEDNVALAVLNDGDPINNTPTVT-GASPGRLRELRLSLTEILKGGVLIQNRNQLCYOD 184
 Db 110 LFEKRYALAVLNDNDPDNVAASPTGRTPEGLRELQLSLTEILKGGVLIQNRNQLCYOD 169
 QY 185 TILMKDIFHKNNQALTLIDTNRSRACHPSMCKSGRCMGESSEDCOSLTRVCAGCA 244
 Db 170 MYLAKDYFRKNNQALPYDIDTNRSRACHPSMCKSGRCMGESSEDCOSLTRVCAGCA 229
 QY 245 RCKRPLPTDCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMNP 304
 Db 230 RCKRPLPTDCHEGCAAGCTGPKHSDCLACHFNHSGICELHCPALVTYNTDFESMNP 289
 QY 305 EGRTPGASCTACPYNYLSTVGS----- 329
 Db 290 EGRTPGASCTACPYNYLSTVGSCTLVCPNNQEVTAEDGTORCEKSPCARVCTGL 349
 QY 330 ----- 329
 Db 350 GMEHLRGARATISDNQGEFDCGCKIFGSLAFLPESFDODPSGIAPRPEQLQVFETLEE 409
 QY 330 ----- 329
 Db 410 ITGYLYISAMPDSLRLDSVFQNLRIINGRLIHGAYSLTLOGIGIHSLGRLSRLSGSL 469
 QY 330 ----- 329
 Db 470 ALIHRNAHLCTVHTVPMDQLFRNRHQAALHSGNRPEEDLCVSSGLVCNSLCAHGHCMGPG 529
 QY 330 ----- 329
 Db 530 PTCVCNCSHLRGOECYEECRVMKGLPREVYSKRLCPHCEQOPNSSETCFGSEADQC 589
 QY 330 ----- 329
 Db 590 AACAHYDSSSCVAPCEGVPDLSTYPMKYRDEEGICOPCPINCTHSCVDLDERGCPA 649
 QY 330 ----- 329
 Db 650 EQRASPTFTIATVEGVLFLILVVVVGILIKRRKQIRKTYMRRLQETELVEPLTPSG 709
 QY 330 ----- 329
 Db 710 AMPNQAQMRILKETELRKVKVLGSGAGTYVKGIMPDGENVKIPVAIKVLRNTSPKAN 769
 QY 330 ----- 329
 Db 770 KEILDEAYVAGVSPYVSRLLIGLITSTVOLYQLMPYGLLDHVEHGRGLSGODLNN 829
 QY 330 ----- 329
 Db 830 WCVOIAGKMSYLEDLVLRHDLAARNVLKSPNHVKTITDEGLARLLDITETEHADGKV 889
 QY 330 ----- 329

Db 890 PIKMALESILRRRFTHQSDVWSYGVWELMPGAKPYDGIAPRIPDLLEKGFRLPDP 949
 QY 330 ----- 329
 Db 950 PICTIDYIMVIMVKCMIDSECRPRFRRELVEFSHMAPDRFVVIQNEDLGSSPYMDTF 1009
 QY 330 ----- 329
 Db 1010 YRSLLEDDMDGLVDAAEYLYPQGGFSPDPTGTGTSTARRHRSSSTRGCGCLTGLLE 1069
 QY 357 PSEEBARSPPLAPSEGAGSDVFDGLMGAAKGLQSLPTHDPSPLQRYSDPTVLPSET 416
 Db 1070 PSEEBARSPPLAPSEGAGSDVFDGLMGAYTKGLQSLPHDLSPQLRSEDPTLPLPET 1129
 QY 417 DGYAPRLTCSQPPRYVQNPDRPQPPSPRGCPRLPAARACATLELATLSRKNQGVKDY 476
 Db 1130 DGYAPRLTCSQPPRYVQNPDRPQPPSPRGCPRLPAARACATLELATLSRKNQGVKDY 1189
 QY 477 FAFGAVENPEYLPPOGGAAPQHPAPFAPFADNLTYMDODPRERGAAPSTFKGPTAE 536
 Db 1190 FAFGAVENPEYLPPOGGAAPQHPAPFAPFADNLTYMDODPRERGAAPSTFKGPTAE 1249
 QY 537 NPEYLGIDVP 546
 Db 1250 NPEYLGIDVP 1259

RESULT 3

148161
 P:185 precursor - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: 148161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishika
 Gene 140, 251-255, 1994
 A:Title: Cloning and activation of the Syrian hamster new proto-oncogene.
 A:Reference number: 148161; MUID:94193007; PMID:7908275
 A:Accession: 148161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1254 <RES>
 A:Cross-references: GB:D16295; NID:g493236; PIDN:BA003801.1; PID:g747595
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif

Query Match 51.84; Score 1935; DB 2; Length 1254;
 Best Local Similarity 33.94; Pred. No. 1.1e-99;
 Matches 423; Conservative 32; Mismatches 66; Indels 728; Gaps 2;

QY 25 WLDKSVLAKELARGASTOVCTGTDMLRLPASPETHIDMLRLHYOGGVVGNLELTYL 84
 Db 6 WCGKGLLALLSPGASSTOVCTGTDMLRLPASPETHIDIVRLHYOGGVVGNLELTYL 65
 QY 85 PTNALSFLQDIQVGVYLAHNOVROYDRLRIVRGTOLEFEDNVALAVLNDGDPDNN 144
 Db 66 PANATLSFLQDIQVGVYLAHNSQVRHVPRLRIVRGTOLEFEDKVALAVLNDRDPDNN 125
 QY 145 TTPYTGASPGRLRELRLSLTEILKGGVLIQNRNQLCYODTILMKDIFHKNNQALTLID 204
 Db 126 VTATGRTPEGLRELQLSLTEILKGGVLIQNRNQLCYODTILMKDIFHKNNQALPYDID 185
 QY 205 TNRSRACHPSMCKSGRCMGESSEDCOSLTRVCAGCAGCARCKRPLPTDCHEGCAAGCT 264
 Db 186 TNRSRACHPSMCKSGRCMGESSEDCOSLTRVCAGCAGCARCKRPLPTDCHEGCAAGCT 245
 QY 265 GPKHSDCLACHFNHSGICELHCPALVTYNTDFESMNPDEGRYTFGASCTACPYNYLS 324
 Db 246 GPKHSDCLACHFNHSGICELHCPALVTYNTDFESMNPDEGRYTFGASCTACPYNYLS 305

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Jun-1999
 C:Accession: A42032
 R:Flickinger, T.W.; Mahle, N.J.; Kung, H.J.
 Mol. Cell Biol. 12, 883-893, 1992
 A:Title: An alternatively processed mRNA from the avian c-erbB gene encodes a soluble, t
 A:Reference number: A42032; MUID:92123214; PMID:1732751
 A:Accession: A42032
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-527 <FL>
 A:Cross-references: GB:M7637; NID:g211737; PIDN:AAA48759.1; PID:g211738
 A:Experimental source: liver
 A:Note: sequence extracted from NCBI backbone (NCBIN:76892, NCBI:76893)
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor

Query Match 20.2%; Score 754.5; DB 2; Length 527;
 Best Local Similarity 42.4%; Pred. No. 1.1e-34;
 Matches 160; Conservative 56; Mismatches 126; Indels 35; Gaps 10;

Oy 19 LELFFWIDRSVLAKELARGASTGVCTGTDKRLRLPASPEHLDMLRHLGYGCGVQGN 78
 18 LVLLLLLRVALCS-----ANEKKVCOGTNNKLTQLGHVEHFSLOMYNNCEVLSN 73
 Oy 79 LELTLPNASTLFDIOEVGYVLIANQVQVPLQRLRIIVRGTOLEFEDNYALAVLDN 138
 74 LEITVEHNRDLTFLKLTQEVAGVYLALNMVDYPLEMLQIIRGNVLDNSFALAVLSN 133
 Oy 139 GDPNNTPTVYGASPGGLRELQRLSLTEILKGVLIQRPOLCYDPTIMKDIPIKNNOL 198
 134 YH-MNKTO-----GLRELPMKRLSELINGVKISNNPKLCNMDVLMNDIIDSRSK- 183
 Oy 199 ALTLID-TNRSRACHPCSPMGSGSRMGSSSDCSLTRTVAGCA-RCRKLPTDCH 256
 184 PLTVLDFPNSLSSCPKCHNCTEDHCWGAGEONCQTLTKVTDACQSGCRKRVSDCH 243
 Oy 257 ECGAAGCTGPKHSDCLACLFHNSICELHCPALVTYNTDFEESNPFGRTFASCVT 316
 244 NCGAAGCTGPRSDCLACKFRDQATKDTCPVLVNTTYQMDVNPBGKTSFQATCVR 303
 Oy 317 ACPYVLTSTDVSGAGGVNHRHRSSTRSGGDLTLGLEPSEEAAPSLAPSEAGSD 376
 304 ECPHYVVTVDHSCV-----RSCNTDT-----YEVEENGVRK-CKKCGLCSK 345
 Oy 377 VFDGDLGMAAGLOSL 393
 346 VCGN-IGIGELKGLI 361
 Db

RESULT 6
 GOHUE
 epidermal growth factor receptor precursor - human
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
 C:Species: Homo sapiens (man)
 C:Date: 15-Nov-1984 #sequence_revision 27-Nov-1985 #text_change 11-Jun-1999
 C:Accession: A00641; A25772; S30024; A38672; A33615; A23062; A05281; A60143; A33
 R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y
 yg, P.H.
 Nature 309, 418-425, 1984
 A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of
 A:Reference number: A00641; MUID:84219729; PMID:6328312
 A:Accession: A00641
 A:Molecule type: mRNA
 A:Residues: 1-1210 <ULL>
 A:Cross-references: EMBL:X00588; NID:g31113; PIDN:CAA23240.1; PID:g757924
 A:Note: the authors translated the codon AAG for residue 540 as Asn
 R:Shih, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.
 Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985
 A:Title: Characterization and sequence of the promoter region of the human epidermal gro
 A:Reference number: A25772; MUID:85270438; PMID:2991899
 A:Accession: A25772
 A:Status: translation not shown
 A:Molecule type: DNA

A:Residues: 1-29 <ISH>
 A:Cross-references: GB:M1234; NID:g181981; PIDN:AAA52370.1; PID:g553272
 R:Haley, J.; Whittle, N.; Bennett, P.; Kitchington, D.; Ulrich, A.; Waterfield, M.
 Oncogene Res. 1, 375-396, 1987
 A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification
 A:Reference number: S30024; MUID:88217333; PMID:3329716
 A:Accession: S30024
 A:Molecule type: DNA
 A:Residues: 1-29 <HA2>
 A:Cross-references: EMBL:X06370; NID:g31118; PIDN:CAA29668.1; PID:g31119
 R:Haley, J.D.; Waterfield, M.D.
 J. Biol. Chem. 266, 1746-1753, 1991
 A:Title: Contributory effects of de Novo transcription and premature transcript termi
 A:Reference number: A38672; MUID:91107677; PMID:1988448
 A:Accession: A38672
 A:Molecule type: DNA
 A:Residues: 1-29 <HAL>
 A:Cross-references: GB:M38425; NID:g181977; PIDN:AAA63171.1; PID:g553271
 A:Experimental source: carcinoma cell line A431-7
 R:Xu, Y.; Shih, S.; Clark, A.J.L.; Sullivan, M.; Wilson, R.K.; Ma, D.P.; Roe, B.A.;
 Nature 309, 806-810, 1984
 A:Title: Human epidermal growth factor receptor cDNA is homologous to a variety of RN
 A:Reference number: A00642; MUID:84245835; PMID:6330563
 A:Accession: A00642
 A:Molecule type: mRNA
 A:Residues: 'RCAMRRA', 150-187, 'KSVIOAV', 195, 'W', 197, 'A', 199-222, 'S', 224-304, 'RA', 307-
 '798-799, 'ND', 802-811, 'R', 813-942 <XUV>
 A:Experimental source: A431 human carcinoma cells, which have large numbers of EGF re
 R:Lin, C.R.; Chen, W.S.; Krulger, W.; Stolarsky, L.S.; Weber, W.; Evans, R.M.; Verma,
 Science 224, 843-848, 1984
 A:Title: Expression cloning of human EGF receptor complementary DNA: gene amplificacl
 A:Reference number: A43615; MUID:84196372; PMID:6326261
 A:Accession: A43615
 A:Molecule type: mRNA
 A:Residues: 713-964 <LIN>
 A:Experimental source: epidermoid carcinoma cell line A431
 R:Slimem, F.A.; Gope, M.L.; Schultz, T.Z.; Wright, D.A.; Carpenter, G.; O'Malley, B.W.
 Biochem. Biophys. Res. Commun. 124, 125-132, 1984
 A:Reference number: A23062; MUID:85046483; PMID:6093780
 A:Accession: A23062
 A:Molecule type: mRNA
 A:Residues: 1028-1210 <SIM>
 R:Weber, W.; Gull, G.N.; Spels, J.
 Science 224, 294-297, 1984
 A:Reference number: A05281; MUID:84172183; PMID:6324343
 A:Accession: A05281
 A:Molecule type: protein
 A:Residues: 25-30, 'S', 32-51, 454-467 <WEB>
 R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.
 J. Biol. Chem. 260, 5205-5208, 1985
 A:Title: Identification of residues in the nucleotide binding site of the epidermal g
 A:Reference number: A60143; MUID:85182650; PMID:298580
 A:Accession: A60143
 A:Molecule type: protein
 A:Residues: 740-744, 'X', 746-747 <RUS>
 R:Morczkowski, B.; Mosig, G.; Cohen, S.
 Nature 309, 270-273, 1984
 A:Title: ATP-stimulated interaction between epidermal growth factor receptor and supe
 A:Reference number: A38023; MUID:84191554; PMID:6329948
 A:Contents: annotation; receptor activity
 A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA
 R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C
 Cell 59, 33-43, 1989
 A:Title: Functional independence of the epidermal growth factor receptor from a domai
 A:Reference number: A3331; MUID:90003233; PMID:2790960
 A:Contents: annotation; internalization signal
 C:Comment: Binding of EGF to the receptor leads to internalization of the EGF-recepto
 C:Genetics:
 A:Gene: GDB:EGFR
 A:Cross-references: GDB:120610; OMIM:131550
 A:Map position: 7p12.3-7p12.1
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos

A:Cross-references: GB:L068664; NID:g193001; PIDN:AAA53029.1; PID:g567201
 C:Genetics:
 A:Gene: EGFR
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphop
 F:1-24/Domain: signal sequence #status predicted <Stic>
 F:648-670/Domain: transmembrane #status predicted <TM>
 F:712-977/Domain: protein kinase homology <KIN>
 F:720-728/Region: protein kinase ATP-binding motif
 F:680-695/Binding site: phosphate (Thr) (covalent) #status experimental
 F:697-1070/Binding site: phosphate (Ser) (covalent) #status experimental
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 19.1%; Score 717.5; DB 2; Length 1210;
 Best Local Similarity 46.8%; Pred. No. 3.1e-32;
 Matches 141; Conservative 39; Mismatches 100; Indels 21; Gaps 5;

36 ARGAA--STOVCTGDMKRLRLPASPTHLDMRLHLYOGCVOGNLELTPTNASTLFL 93
 20 AAGGALEKKVVCOSTNRLLTOLGFEDHFLSLDRMYNCEVVLGNLETTVYQNNYDSEL 79
 94 QDIOEVGYVLIANOVROYPLQRLIRVGTODLFEDNYALAVLNDGDPNLNTPVTCASP 153
 80 KTIQEVAGYVLIANTVERIPLENLQIRGNALYENTYALALISN-----YGTNR 129
 154 GGLREQLRSTLEILKGVLIQIRNPOLCYODTILMKDI----FHKNNQLALTILDRNR 209
 130 TGRLEPLMRNLQILIGAVFSSNNPLICNMDTQMRIDYONVMSNMSML----OSHPS 185
 210 ACHPCSPMKSGRSGWSESSDCSLRTVYAGGCA--RCKGPLPTDCHEQCAAGCTGPKH 268
 186 SCRCPCPSGNGSCWGGEENCKLTKIICAOQCSHRCSRSPSDCHNCNCAAGCTGPRE 245
 269 SDCGLACIHRHSGICELHCPALVTYNTDFESMPNPGRTFGASCYTACPYNLSTDVG 328
 246 SDCLVCKCFODEATCKDKTCLPLMLYNPTTYQMDVNPEGKYSFGATCVKCKPRNVVVDHG 305
 329 S 329
 306 S 306

RESULT 9
 A36325
 epidermal growth factor receptor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 10-Oct-1997
 C:Accession: A36325
 R:Petich, L.A.; Harris, J.; Raymond, V.W.; Blasband, A.; Lee, D.C.; Earp, H.S.
 Mol. Cell. Biol. 10, 2973-2987, 1990
 A:Title: A truncated, secreted form of the epidermal growth factor receptor is encoded b
 A:Reference number: A36325; MUID:90258888; PMID:2342466
 A:Accession: A36325
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-644 <PEP>
 A:Cross-references: GB:M37394
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: alternative splicing; ATP; growth factor receptor

Query Match 19.1%; Score 714.5; DB 2; Length 644;
 Best Local Similarity 37.2%; Pred. No. 3.e-32;
 Matches 169; Conservative 60; Mismatches 158; Indels 67; Gaps 13;

30 VLAKELARGAA--STOVCTGDMKRLRLPASPTHLDMRLHLYOGCVOGNLELTPTN 87
 14 LLAALCAAGALBEKKVVCOSTNRLLTOLGFEDHFLSLDRMYNCEVVLGNLETTVYQNR 73
 88 ASLSFLDIOEVGYVLIANOVROYPLQRLIRVGTODLFEDNYALAVLNDGDPNLNTP 147
 74 YDLSFLTKTIOEVAGYVLIANTVERIPLENLQIRGNALYENTYALALISN----- 124

148 VTGASPGGLREQLRSTLEILKGVLIQIRNPOLCYODTILMKDIFHKNNQLALTILDTNR 207
 125 -YGTNNKTGRLRLMRNLQILIGAVFSSNNPLICNMDTQMRIDYONVMSNMSMLDYOR 182
 208 S-RACHPCSPMKSGRSGWSESSDCSLRTVYAGGCA--RCKGPLPTDCHEQCAAGCTG 265
 183 HLNGCPKCDPSCPNCSWGGEENCKLTKIICAOQCSHRCSRSPSDCHNCNCAAGCTG 242
 266 PKHSDDLACIHRHSGICELHCPALVTYNTDFESMPNPGRTFGASCYTACPYNLST 325
 243 PRSDCLVCKCFODEATCKDKTCLPLMLYNPTTYQMDVNPEGKYSFGATCVKCKPRNVV 302
 326 DVCSGAGGWNHHRSSSTRSGGDLTLGLPESEEARPLAPSGCAGSDVFDGDLGMC 385
 303 DHG-----SCVACCPDIY-----EVEEDGYSKCKKCDGPKKVCNG-IGIG 343
 386 AAKGLQSLPTHPDPSPLQRYSEDPVPLPSETDGVAPLTCSPQPEVYNQDPVRPQPSPR 445
 344 EFRDITLST---NMDNIRKHFYCTAI-----SGDLHLPLVA-----FKGDSFTTRPPLDPR 390
 446 EGPLPAARPAATLERAKTILSPGKNGVYKDFEAF 479
 391 E-----LEILKT-----VKEITGF 404

RESULT 10
 A36223
 kinase-related transforming protein (erbB3) (EC 2.7.1.-) precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 04-Oct-1991 #sequence_revision 13-Jan-1993 #text_change 17-Nov-2000
 C:Accession: A36223; 159164
 R:Raus, M.H.; Issing, W.; Maki, T.; Popescu, N.C.; Aaronson, S.A.
 Proc. Natl. Acad. Sci. U.S.A. 86, 9193-9197, 1989
 A:Title: Isolation and characterization of ERBB3, a third member of the ERBB/epiderma
 A:Reference number: A36223; MUID:90083234; PMID:2687875
 A:Accession: A36223
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1342 <KRA>
 A:Cross-references: GB:M29366
 R:Plowman, G.D.; Whitney, G.S.; Neubauer, M.G.; Green, J.M.; McDonald, V.L.; Todaro,
 Proc. Natl. Acad. Sci. U.S.A. 87, 4905-4909, 1990
 A:Title: Molecular cloning and expression of another epidermal growth factor receptor
 A:Reference number: 159164; MUID:90311312; PMID:2164210
 A:Accession: 159164
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-559, 'G', 561-957, 'F', 959-1063, 'G', 1065-1342 <RES>
 A:Cross-references: GB:M34309; NID:g183990; PIDN:AAA53979.1; PID:g106841
 C:Genetics:
 A:Gene: GDB:ERBB3; HER3
 A:Cross-references: GDB:119880; OMIM:190151
 A:Map position: 12q13-12q13
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; phosphotransferase
 F:707-972/Domain: protein kinase homology <KIN>
 F:715-723/Region: protein kinase ATP-binding motif

Query Match 18.6%; Score 697; DB 2; Length 1342;
 Best Local Similarity 44.6%; Pred. No. 4.8e-21;
 Matches 139; Conservative 45; Mismatches 98; Indels 30; Gaps 7;

16 LGFLFLFFFLDLRSVLAKELARGAA--STOVCTGDMKRLRLPASPTHLDMRLHLYOGC 72
 10 LGILF-----STARSGEVSQAVCPGTNLGSLVGDAQNVQYQTLKLTERC 56
 73 QVQGNLELTPTNASTLSFLDIOEVGYVLIANOVROYPLQRLIRVGTODLFEDNYA 132
 57 EYVMGNLEIYLTGNADSLQWIREVTGYVLAMNFSFLPLPNLRVAVGQVYDGKFA 116
 133 LAVLNDGDPNLNTPPYTGASPGGLREQLRSTLEILKGVLIQIRNPOLCYODTILMKDIF 192

Db 117 IFPM-----LNYNT-----NSSHALRLQLTQLTEILISGVIYEKNDLCHMDITDMRDIV 167
 Qy 193 HKNQNALTLIDNRFRACHPCSPMCKSGRSGESSEDSCSLTRTYCAGGC-ARCKGPLP 251
 Db 168 RDRD---AEIVYKNDGNSCPCHVEYCKG-RCWGPGSEDCQTLTKRTICAPCNCHGCEPNP 223
 Qy 252 TDCHECAGACGCPKHSQDLACLHFNHSGICELHCPALVTYNTDFESMPNPGRTYG 311
 Db 224 NQCHHECAGACGCPQDPTDFACRHNDSGACVPCRPQPLVYKNTLFELEPNPHTKYQYG 283
 Qy 312 ASCTACPCNYUL 323
 Db 284 GVCVASCPNPFV 295

RESULT 11
 FOHUGM
 granulocyte-macrophage colony-stimulating factor precursor (validated) - human
 N:Alternate names: colony-stimulating factor 2; GM-CSF
 C:Species: Homo sapiens (man)
 C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 08-Dec-2000
 C:Accession: C24636; 159065; A25169; A01853; A44175; JCI090
 R:Myatake, S.; Otsuka, T.; Yokota, T.; Lee, F.; Aral, K.
 EMBL J. 4. 2561-2568, 1985
 A>Title: Structure of the chromosomal gene for granulocyte-macrophage colony stimulating
 A:Reference number: A91015; MUID:86030234; PMID:3876930
 A:Accession: C24636
 A:Molecule type: DNA
 A:Residues: 1-144 <MIV>
 A:Cross-references: EMBL:X03021; NID:g31858; PIDN:CAA26822.1; PID:g31859
 R:Kaushansky, K.; O'Hara, P.J.; Berkner, K.; Segal, G.M.; Hagen, F.S.; Adamson, J.W.
 Proc. Natl. Acad. Sci. U.S.A. 83, 3101-3105, 1986
 A>Title: Genomic cloning, characterization, and multilineage growth-promoting activity
 A:Reference number: 159065; MUID:86205844; PMID:3486413
 A:Accession: 159065
 A>Status: translated from GB/EMBL/DBDUT
 A:Molecule type: DNA
 A:Residues: 1-144 <RES>
 A:Cross-references: GB:M13207; NID:g181147; PIDN:AAA98768.1; PID:g181148
 R:Cantrill, M.A.; Anderson, D.; Cerretti, D.P.; Price, V.; McKereghan, K.; Tushinski, R.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6250-6254, 1985
 A>Title: Cloning, sequence, and expression of a human granulocyte/macrophage colony-stim
 A:Reference number: A25169; MUID:8528329; PMID:3898082
 A:Accession: A25169
 A:Molecule type: mRNA
 A:Residues: 1-144 <CAN>
 A:Cross-references: GB:M1734; NID:g181149; PIDN:AAA52122.1; PID:g181150
 R:Lee, F.; Yokota, T.; Otsuka, T.; Gemmell, L.; Larson, N.; Luh, J.; Aral, K.; Rennick,
 Proc. Natl. Acad. Sci. U.S.A. 82, 4360-4364, 1985
 A>Title: Isolation of cDNA for a human granulocyte-macrophage colony-stimulating factor
 A:Reference number: A01853; MUID:85242684; PMID:3925454
 A:Accession: A01853
 A:Molecule type: mRNA
 A:Residues: 1-144 <LEES>
 A:Cross-references: GB:M11220; NID:g183363; PIDN:AAA52578.1; PID:g183364
 R:Hong, G.G.; Witek, J.S.; Temple, P.A.; Wilkens, K.M.; Leary, A.C.; Luxenberg, D.P.; Jo
 A.: Clark, S.C.
 Science 248, 810-815, 1985
 A>Title: Human GM-CSF: molecular cloning of the complementary DNA and purification of th
 A:Reference number: A44175; MUID:85218749; PMID:3923623
 A:Accession: A44175
 A:Molecule type: mRNA
 A:Residues: 1-116, 'T', 118-144 <MON>
 A:Cross-references: GB:M10663; NID:g181145; PIDN:AAA52121.1; PID:g181146
 A:Note: parts of this sequence, including the amino end of the mature protein, were conf
 Acta Biochim. Biophys. Sin. 25, 651-655, 1993
 R:Wen, D.Y.; Huang, B.R.; Cai, L.W.; Si, J.Y.
 A>Title: Amplification of human granulocyte-macrophage colony-stimulating factor cDNA wi
 A:Accession: JCI090
 A:Molecule type: protein
 A:Residues: 18-21, 'C', 23-96, 'L', 98-144 <MEN>
 C:Genetics:

A:Gene: GDB:CSF2
 A:Cross-references: GDB:119812; OMIM:138960
 A:Map position: 5q23.2-5q31.1
 A:Introns: 53/3; 67/3; 109/3
 C:Function:
 A:Description: stimulates the differentiation and proliferation of hematopoietic prog
 C:Superfamily: granulocyte-macrophage colony-stimulating factor
 C:Keywords: cytokine; glycoprotein; growth factor; macrophage; monomer; T-cell
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-144/Product: granulocyte-macrophage colony-stimulating factor #status experiment
 F:44,54/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.08; Score 674; DB 1; Length 144;
 Best Local Similarity 99.28; Pred. No. 7, 5e-31;
 Matches 127; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 548 AAPARSPSPQPEHNAIOEARRLNLSRDAEKNETVEYISEMFDQEPICLOTRL 607
 Db 17 SAPARSPSPQPEHNAIOEARRLNLSRDAEKNETVEYISEMFDQEPICLOTRL 76
 Qy 608 ELXKGLRGSILTKLGPDTMASHYKQCPPTPETSQATQITFESEKKNLKDPLVLPF 667
 Db 77 ELXKGLRGSILTKLGPDTMASHYKQCPPTPETSQATQITFESEKKNLKDPLVLPF 136
 Qy 668 DCWEPEVOE 675
 Db 137 DCWEPEVOE 144

RESULT 12
 JCI4387
 Epidermal growth factor receptor homolog precursor - rat
 N:Alternate names: ErbB3 protein; HER3 protein
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 13-Nov-1998
 C:Accession: JCI4387
 R:Hellier, N.J.; Kim, H.H.; Greaves, C.H.; Sterke, S.L.; Koland, J.G.
 Gene 165, 279-284, 1995
 A>Title: Cloning of the rat ErbB3 cDNA and characterization of the recombinant protei
 A:Reference number: JCI4387; MUID:96096535; PMID:8522190
 A:Accession: JCI4387
 A:Molecule type: mRNA
 A:Residues: 1-1339 <HELD>
 A:Cross-references: GB:U29339; NID:g915389; PID:g915390
 A:Experimental source: liver
 A:Note: The authors translated the codon AAC for residue 369 as Thr and GAT for resid
 C:Comment: This protein is a functional heregulin receptor that transduces signals to
 C:Genetics:
 A:Gene: ErbB3
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; growth factor receptor; liver; phosphoprotein; transmembrane protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1339/Product: epidermal growth factor homolog #status predicted <MAT>
 F:640-659/Domain: transmembrane #status predicted <TM>
 F:705-970/Domain: protein kinase homology <KIN>
 F:713-721/Region: protein kinase ATP-binding motif
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr)
 F:939,1051,1156,1194,1196,1219,1257,1259,1273,1286,1325/Binding site: phosphate (Tyr)

Query Match 17.7%; Score 662; DB 2; Length 1339;
 Best Local Similarity 43.48; Pred. No. 4, 2e-29;
 Matches 135; Conservative 44; Mismatches 108; Indels 24; Gaps 7;

Qy 17 GFLFLFFMIDRSVLAKELARGAA---STQVCTGTDKLRLPASPETHLMDLRLKVGCG 73
 Db 5 GTGIVLCFL-----SLARGSEMGSQAVCGTLNGLSVTDGADNOGYTLKVEKCE 57
 Qy 74 VVGNGLETLPLPNASISFLQDIOEVGYLVLIHNOVROPLORLRYRGTQFEDNYAL 133
 Db 58 VVGNGLEIYLTGNADLSFLQWIREYAYLVAMNESVYLPNLRVRYRGTQYVDGFAI 117
 Qy 134 AVLQNDGPIUNNTTPVVGSGPGRLRLQLRLTEILKGVLIQRPOLCYQDTILMKDIFH 193
 Db 118 FVM-----LNYNT-----NSSHALRLQLTQLTEILISGVIYEKNDLCHMDITDMRDIV 168

Oy 668 DCWEPOE 675
| | | | |
Db 137 DCWEPAOK 144

Search completed: April 28, 2003, 13:42:22
Job time : 32.3846 secs

DR WPI: 2001-662965/76.
 DR N-PSDB: AAD21565.
 PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 PS Claim 7: Page 26; 59pp; English.
 CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC hGM-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC intracellular domain, a mature HER-2 membrane distal extracellular and
 CC macrophage colony stimulating factor (GM-CSF) sequence and a
 CC C-terminal tag.
 XX

SO Sequence 690 AA:

Query Match 100.0%; Score 3739; DB 22; Length 690;
 Best Local Similarity 100.0%; Pred. No. 4.8e-226;
 Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRAAPLLAARASLSGFLFLEFMDRSVLAKELARGAASVGVCTGDMKRLPASPET 60
 DB 1 MRAAPLLAARASLSGFLFLEFMDRSVLAKELARGAASVGVCTGDMKRLPASPET 60
 OY 61 HLDMLRHLVGGCGVGVGNLELTPTNASLSFQDIQEVGVYLIANQVQVPLQRLRI 120
 DB 61 HLDMLRHLVGGCGVGVGNLELTPTNASLSFQDIQEVGVYLIANQVQVPLQRLRI 120
 OY 121 VRCSTQLFEDNYALAVLDNGDPLNNTTPTVGASPGSLREIQLSLTELKGVLIQRPOL 180
 DB 121 VRCSTQLFEDNYALAVLDNGDPLNNTTPTVGASPGSLREIQLSLTELKGVLIQRPOL 180
 OY 181 CVQDTLMDIFKNNQALTLTDNRSRACHSPCKGSRGWGSESSCCSLTRYVCA 240
 DB 181 CVQDTLMDIFKNNQALTLTDNRSRACHSPCKGSRGWGSESSCCSLTRYVCA 240
 OY 241 GCGARCKGRLPTDCHEOCAGCTGPKHSDCLCLFPHNSGICELCPALVTNTDFES 300
 DB 241 GCGARCKGRLPTDCHEOCAGCTGPKHSDCLCLFPHNSGICELCPALVTNTDFES 300
 OY 301 MPNPEGRTYFGASCVTACRYNYLSTDVSGAGGMVHNRHRSSTRSGGDLTLGLEPSEE 360
 DB 301 MPNPEGRTYFGASCVTACRYNYLSTDVSGAGGMVHNRHRSSTRSGGDLTLGLEPSEE 360
 OY 361 EAPRSPLASSEGASGVDFGDLGMAKGLQSLPTHDPSFLQRYSEDPPVPLPSETDGYV 420
 DB 361 EAPRSPLASSEGASGVDFGDLGMAKGLQSLPTHDPSFLQRYSEDPPVPLPSETDGYV 420
 OY 421 APLTCGPOPEYVQNPVROPSPREGPLPAARPAAGATLERATKLSPGKGVKDVFAFG 480
 DB 421 APLTCGPOPEYVQNPVROPSPREGPLPAARPAAGATLERATKLSPGKGVKDVFAFG 480
 OY 481 GAVENBEYLTPOGGAAPQHPAPAFSPADNLYWDDPDERGAPSTEGTPTAENPEY 540
 DB 481 GAVENBEYLTPOGGAAPQHPAPAFSPADNLYWDDPDERGAPSTEGTPTAENPEY 540
 OY 541 LGIDVPAAPARSPSPSTQPMWEHVAIOEARLLNLSRDTAEMNETVEYISMFDLQEP 600
 DB 541 LGIDVPAAPARSPSPSTQPMWEHVAIOEARLLNLSRDTAEMNETVEYISMFDLQEP 600
 OY 601 TCIQTLELYKQGLRSLTKLGLTMMASHYKQHCPPPTSCATQIITFESEKFNMD 660
 DB 601 TCIQTLELYKQGLRSLTKLGLTMMASHYKQHCPPPTSCATQIITFESEKFNMD 660

OY 661 FLIVIPDCWEVQEGARPPRAAANNHNNH 690
 DB 661 FLIVIPDCWEVQEGARPPRAAANNHNNH 690

RESULT 2

AAE13111
 ID AAE13111 standard; Protein; 697 AA.

AC AAE13111;

DT 28-JAN-2002 (first entry)

DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide:

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW Immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
 KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.

XX Chimeric - Homo sapiens.

OS Chimeric - Rattus norvegicus.

PN WO200174855-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US10515.

PR 30-MAR-2000; 2000US-193504P.

PA (DEND-) DENDREON CORP.

PI Laus R, Vidovic D, Graddis T;

DR WPI: 2001-662965/76.

DR N-PSDB: AAD21567.

PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -

PS Claim 7: Page 27; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC rGM-CSF fusion protein construct which comprises human PAP
 CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
 CC signal sequence, mature HER-2 membrane distal extracellular domain,
 CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
 CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
 CC and a C-terminal tag.
 XX

SO Sequence 697 AA:

Query Match 92.9%; Score 3473.5; DB 22; Length 697;
 Best Local Similarity 91.7%; Pred. No. 2.1e-209;
 Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;

OY 1 MRAAPLLAARASLSGFLFLEFMDRSVLAKELARGAASVGVCTGDMKRLPASPET 60
 DB 1 MRAAPLLAARASLSGFLFLEFMDRSVLAKELARGAASVGVCTGDMKRLPASPET 60

Db 1 MRAAPLLARAASLSLGLFLFLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60
QY 61 HLDMLRHLYOGGOVVGMLLELYLPTNASLFLDIOEVGVYLLAHNOVQVPLQRLRI 120
Db 61 HLDMLRHLYOGGOVVGMLLELYLPTNASLFLDIOEVGVYLLAHNOVQVPLQRLRI 120
QY 121 VRGTOLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELOLRSLTEILKGGVLIORNPQL 180
Db 121 VRGTOLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELOLRSLTEILKGGVLIORNPQL 180
QY 181 CYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKGSGRCWSESSEDCOSLTRTYCA 240
Db 181 CYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKGSGRCWSESSEDCOSLTRTYCA 240
QY 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLAHFNHSGICELHCALVTYNTDTRES 300
Db 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLAHFNHSGICELHCALVTYNTDTRES 300
QY 301 MPNEGRTYFGASCYTACRYNYLSTDVGS-----GAGGMVHNHRRSSSTRSGGDL 351
Db 301 MPNEGRTYFGASCYTACRYNYLSTDVGSASINFEKLAGAGMVHNHRRSSSTRSGGDL 360
QY 352 TLGLEPSEEEARSRPLAPSEAGSDVFDGDLGMAKGLQSLPHNDSPRLQRYSEDPYR 411
Db 351 TLGLEPSEEEARSRPLAPSEAGSDVFDGDLGMAKGLQSLPHNDSPRLQRYSEDPYR 420
QY 412 LPSETDGVAVPLTCSPOREYVNOPVROPSPREGRLPAARPAATLERAKTTLSPGKNG 471
Db 421 LPSETDGVAVPLTCSPOREYVNOPVROPSPREGRLPAARPAATLERAKTTLSPGKNG 480
QY 472 VVKDVFARGAVENPEYLTPOGGAARQPHRRPAPSFADNLYWDDPERGAPSTPKG 531
Db 481 VVKDVFARGAVENPEYLTPOGGAARQPHRRPAPSFADNLYWDDPERGAPSTPKG 540
QY 532 TPTAENPEYLGIDVPAARSPSPSTQPMENHVAIOEARRLINLSRPTAAEMNETVEYI 591
Db 541 TPTAENPEYLGIDVPAARSPSPSTQPMENHVAIOEARRLINLSRPTAAEMNETVEYI 600
QY 592 SEMFLOPRTCLQRTLELYKQGLKSLTKLGRPLFMASHYKONRPRPETSQAQIITF 651
Db 601 SNEPSIQARTCVQTRKLKYLKQGLKSLTKLGMALTMASHYQTONCPRLPETSCELEVTTF 660
QY 652 ESFENLKDPLLYIPFDCMERVOEGARRPRAAHNNHNNH 690
Db 661 EDFIKNLGFLFDYRFDGCMKRYQKARPR--ANNHNNH 697

RESULT 3
AAE13108
ID AAE13108 standard; Protein: 555 AA.
XX
AC AAE13108;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct.
XX
KW Immunostimulatory fusion protein; IFP: antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human;
KW HER-2 protein; HER500 fusion protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Synthetic.
XX
PN W0200174855-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US10515.
XX

PR 30-MAR-2000; 2000US-193504P.
XX
PA (DEND-) DENDREON CORP.
XX
PI Laus R, Vidovic D, Graddis T;
XX
PT WPI: 2001-662965/76.
DR N-PSDB: AAD21564.
XX
PT An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and
PT is useful for the treatment of associated cancer associated -
XX
PS
XX
XX Claim 7; Page 26; 59pp; English.
XX
XX The invention relates to immunostimulatory fusion proteins (IFP) and
XX nucleic acid molecules encoding such proteins. The IFPs comprise a
XX polypeptide antigen component and an immunostimulatory component derived
XX from the intracellular domain of HER-2 protein which is effective to
XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX immune response to the antigen. IFP or superactivated dendritic cells
XX are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX associated with a particularly antigen. The present sequence is HER500
XX fusion protein construct which comprises human PAP signal
XX sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
XX sequence, mature HER-2 membrane distal extracellular and intracellular
XX domains and a C-terminal tag.
XX
XX Sequence 555 AA:
XX
XX
XX Query Match 79.5%; Score 2972; DB 22; Length 555;
XX Best Local Similarity 100.0%; Pred. No. 4.1e-178;
XX Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRAAPLLARAASLSLGLFLFLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60
Db 1 MRAAPLLARAASLSLGLFLFLFWLDRSVLAKELARGAASSTOYCTGDMKRLRPASPET 60
QY 61 HLDMLRHLYOGGOVVGMLLELYLPTNASLFLDIOEVGVYLLAHNOVQVPLQRLRI 120
Db 61 HLDMLRHLYOGGOVVGMLLELYLPTNASLFLDIOEVGVYLLAHNOVQVPLQRLRI 120
QY 121 VRGTOLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELOLRSLTEILKGGVLIORNPQL 180
Db 121 VRGTOLFEDNVALAVLDNGDPLNNTPTVTGASPGGLRELOLRSLTEILKGGVLIORNPQL 180
QY 181 CYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKGSGRCWSESSEDCOSLTRTYCA 240
Db 181 CYODTILMKDIFHKNNQALATLIDTNRSRACHPCSPMKGSGRCWSESSEDCOSLTRTYCA 240
QY 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLAHFNHSGICELHCALVTYNTDTRES 300
Db 241 GGCARCKGRLPTDCHEOCAGCTGPKHSDCLAHFNHSGICELHCALVTYNTDTRES 300
QY 301 MPNEGRTYFGASCYTACRYNYLSTDVGSAGAGMVHNHRRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNEGRTYFGASCYTACRYNYLSTDVGSAGAGMVHNHRRSSSTRSGGDLTLGLEPSEE 360
QY 361 EAPSRPLAPSEAGSDVFDGDLGMAKGLQSLPHNDSPRLQRYSEDPYRPLPSETDGYV 420
Db 361 EAPSRPLAPSEAGSDVFDGDLGMAKGLQSLPHNDSPRLQRYSEDPYRPLPSETDGYV 420
QY 421 APRTCSPOREYVNOPVROPSPREGRLPAARPAATLERAKTTLSPGKNGVAVKVPFAG 480
Db 421 APRTCSPOREYVNOPVROPSPREGRLPAARPAATLERAKTTLSPGKNGVAVKVPFAG 480
QY 481 GAVENPEYLTPOGGAARQPHRRPAPSFADNLYWDDPERGAPSTPKGPTAENPEY 540
Db 481 GAVENPEYLTPOGGAARQPHRRPAPSFADNLYWDDPERGAPSTPKGPTAENPEY 540
QY 541 LGIDVPAAR 549
Db 541 LGIDVPAAR 549

RESULT 4
AAE13110
ID AAE13110 standard; Protein; 564 AA.
XX
XX AAE13110:
XX
XX 28-JAN-2002 (first entry)
XX
XX Human HER500 fusion protein construct comprising OVA-derived octapeptide.
DE
XX
XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
KM Immunostimulatory component; T-cell mediated immune response; DC;
KM dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KM PAP protein; Ala Arg linker; membrane distal extracellular domain;
KM membrane distal intracellular domain; C-terminal tag; human; OVA;
KM HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
XX MO200174855-A2..
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US10515.
XX
XX 30-MAR-2000; 2000US-193504P.
XX
XX (DEND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddis T;
PI
XX
XX WPI: 2001-662965/76.
DR N-PSDB; AAD21566.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and
PT is useful for the treatment of associated cancer associated -
XX
XX
PS Claim 7; Page 26; 59pp; English.
XX
XX The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFP or superactivated dendritic cells
CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC sequence, mature HER-2 membrane distal extracellular domain, an
CC Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
CC HER-2 membrane distal intracellular domain and a C-terminal tag.
XX
XX Sequence 564 AA:
SO

Query Match 79.1%; Score 2957.5; DB 22; Length 564;
Best Local Similarity 98.4%; Pred. No. 3.4e-177;
Matches 349; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 MRAPLLAARASISLGLFLFEFWLDRSVLAKEELARGAASSTGVCTGDKMLRLPASPER 60
DB 1 MRAPLLAARASISLGLFLFEFWLDRSVLAKEELARGAASSTGVCTGDKMLRLPASPER 60
QY 61 HLDMLRLHYGGCCVGVGNELELYLPTNASISFLQDIOGVGVYLIANOVROYPLQRLRI 120
DB 61 HLDMLRLHYGGCCVGVGNELELYLPTNASISFLQDIOGVGVYLIANOVROYPLQRLRI 120
QY 121 VRGTQLFEDNRYALAVLDNGDP LNNTPTVTCASPGCLRELDLRSTLTCLKGVLIQRNPOL 180
DB 121 VRGTQLFEDNRYALAVLDNGDP LNNTPTVTCASPGCLRELDLRSTLTCLKGVLIQRNPOL 180

DB 121 VRGTQLFEDNRYALAVLDNGDP LNNTPTVTCASPGCLRELDLRSTLTCLKGVLIQRNPOL 180
QY 181 CYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPMCKSGRCGSESEDCQSLTRTVCA 240
DB 181 CYODTILMKDIFHKNNOLATLIDTNRSRACHPCSPMCKSGRCGSESEDCQSLTRTVCA 240
QY 241 GGCARCKGPLEPTDCCHQCAAGCTGPRHSDCLACLFHNHSGICELCPALVYNTDFEES 300
DB 241 GGCARCKGPLEPTDCCHQCAAGCTGPRHSDCLACLFHNHSGICELCPALVYNTDFEES 300
QY 301 MPNPEGRITFGASCVTACPNYVSTDVGS-----GAGGWHHHHRSSTSGGDL 351
DB 301 MPNPEGRITFGASCVTACPNYVSTDVGS-----GAGGWHHHHRSSTSGGDL 351
QY 352 TLGLEPEEERAPRPLAPSEAGSDVDGDLGMAKGLSLPTPHDPSPLQRYSEDPVP 411
DB 352 TLGLEPEEERAPRPLAPSEAGSDVDGDLGMAKGLSLPTPHDPSPLQRYSEDPVP 411
QY 361 TLGLEPEEERAPRPLAPSEAGSDVDGDLGMAKGLSLPTPHDPSPLQRYSEDPVP 420
DB 361 TLGLEPEEERAPRPLAPSEAGSDVDGDLGMAKGLSLPTPHDPSPLQRYSEDPVP 420
QY 412 LPSETDGYVAPLTCSPQPEYVNDPVRPQPPSPREGPLPAARPAATLERAKTILSPCKNG 471
DB 412 LPSETDGYVAPLTCSPQPEYVNDPVRPQPPSPREGPLPAARPAATLERAKTILSPCKNG 471
QY 472 VVKDVFAGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYWDDPPERGAAPSTFKG 531
DB 472 VVKDVFAGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYWDDPPERGAAPSTFKG 531
QY 481 VVKDVFAGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYWDDPPERGAAPSTFKG 540
DB 481 VVKDVFAGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYWDDPPERGAAPSTFKG 540
QY 532 TPTAENPEYLGIDVPA 549
DB 541 TPTAENPEYLGIDVPA 558

RESULT 5
ID AAB21203
XX AAB21203 standard; protein; 919 AA.
XX
XX AAB21203:
XX
XX 12-JAN-2001 (first entry)
XX
XX Human HER-2/neu fusion protein.
DE
XX
XX Human; HER-2/neu; oncogene; tyrosine kinase; cytostatic; vaccine;
KM breast cancer; prostate cancer; ovarian cancer; lung cancer;
KM colon cancer; fusion protein.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO200044899-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000MO-US02164.
XX
XX 29-JAN-1999; 99US-0117976.
XX
XX (CORI-) CORIXA CORP.
XX (SMIR) SMITHKLINE BEECHAM.
XX
XX Cheever MA, Cheysen D;
PI
XX
XX WPI: 2000-505976/45.
XX
XX HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT useful for vaccinating against breast, ovarian, colon, lung and
PT prostate cancers -
XX
XX Claim 2; Fig 12; 128pp; English.
XX
XX The present sequence is a fusion protein comprising the extracellular
CC domain and the phosphorylation domain of the human HER-2/neu protein.
CC HER-2/neu is a member of the tyrosine kinase family of receptor-like
CC glycoproteins and shows homology to the epidermal growth factor receptor

CC (EGFR). It probably plays a part in cell growth and/or differentiation.
 CC The HER-2/neu gene is an oncogene. HER-2/neu fusion proteins may be used
 CC to treat or prevent cancer by eliciting or enhancing an immune response
 CC to the HER-2/neu protein. They may be used to treat malignancies such as
 CC breast, ovarian, colon, lung and prostate cancers, and may be used as an
 CC antigen to vaccinate against these neoplasias.

XX Sequence 919 AA;

Query Match 68.8%; Score 2573.5; DB 21; Length 919;

Best Local Similarity 56.4%; Pred. No. 7,1e-153;

Matches 509; Conservative 0; Mismatches 3; Indels 391; Gaps 1;

```

OY 35 LARGAASVTCGTDMKRLPASPETHLMDLRHLYGCGVQVGNLELYLPTNASLSFQ 94
DB 16 LPPGAASVTCGTDMKRLPASPETHLMDLRHLYGCGVQVGNLELYLPTNASLSFQ 75
OY 95 DIOEVQGYVLIANQVQVPLQRLRIYRGTLFEDNYALAVLNDGDPINNTPTVYGASPG 154
DB 76 DIOEVQGYVLIANQVQVPLQRLRIYRGTLFEDNYALAVLNDGDPINNTPTVYGASPG 135
OY 155 GLRELQLSLLEILKGGVLIQNRNQLCYQDTILMKDIFHKNNQALATLIDNRSRACHPC 214
DB 136 GLRELQLSLLEILKGGVLIQNRNQLCYQDTILMKDIFHKNNQALATLIDNRSRACHPC 195
OY 215 SPMCKSGKMGESSEDCOSLTRVCAGGACARCKGPLPTDCHECCAGCTGPKHSDCLAC 274
DB 196 SPMCKSGKMGESSEDCOSLTRVCAGGACARCKGPLPTDCHECCAGCTGPKHSDCLAC 255
OY 275 LHFHNSGICELHCPALVYNTDTEESMPNPGRYTFGASCYACPYNTLSTDVGS----- 329
DB 256 LHFHNSGICELHCPALVYNTDTEESMPNPGRYTFGASCYACPYNTLSTDVGSCTLVC 315
OY 330 ----- 329
DB 316 PLHNOVTAEADGTORCEKSPCARVCYGLMEHLREYRAVTSANIOEAGCKKIFGSLA 375
OY 330 ----- 329
DB 376 FLPSFDGDPASNTAPLQPELOVFELEETGYLYISAMPDLSLPDLSPQNLQYIRRI 435
OY 330 ----- 329
DB 436 LHNAGASLTLLGLISWLGSLRLSRLSGSLALIHNNHLCFVNHVPMQDLFRNPHQALLH 495
OY 330 ----- 329
DB 496 TANRPEDECVGEGLACHQLCARHGCMGPGPTQCVCNCSQFLNGQCEVCRLVQLPREYV 555
OY 330 ----- 329
DB 556 NARHCLPCHPECPQNSVTCFGEADQCVACAHKDPFCVCARCPGVKDLISYPIWK 615
OY 330 ----- 329
DB 616 FPDEGACQPCPINCTHSCVDLDKGCFAEORASPLTSQNELDGPAFLDSTFYRSLLED 675
OY 330 ----- 329
DB 676 DDMKDLVADEEYLVPPQGFECPRDPAFGAGVHHRRHSSSTRSGGDLTLLEPSEEBAP 735
OY 364 RSLPLASEGASDVDFDGLGWAAGKGLQSLPTNHPSPLOQRYSEDPYVLPSETDGYVAPL 423
DB 736 RSLPLASEGASDVDFDGLGWAAGKGLQSLPTNHPSPLOQRYSEDPYVLPSETDGYVAPL 795
OY 424 TCSQPEYVNPQVPPSPRREGPLPAARPAATLERAKTLSPGKGVNVDAFAGAV 483
DB 796 TCSQPEYVNPQVPPSPRREGPLPAARPAATLERAKTLSPGKGVNVDAFAGAV 855
OY 484 ENPEYTLPPGGGAAPQHPHPPAFSPAFNLLYWDPPRGRGAPSTFKTPTAENPEYIGL 543
DB 856 ENPEYTLPPGGGAAPQHPHPPAFSPAFNLLYWDPPRGRGAPSTFKTPTAENPEYIGL 915

```

OY 544 DVP 546
 DB 916 DVP 918

RESULT 6

ID AAM51148

AC AAM51148 standard; Protein: 919 AA.

AC AAM51148;

DT 17-JUN-2002 (first entry)

DE Her-2/neu extracellular domain-phosphorylation domain fusion.

KW Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;

KW Tyrosine kinase; receptor; c-erbB2; gene therapy.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Domain 1..653

FT Domain /note="extracellular domain"

FT Domain 654..919

FT Domain /note="phosphorylation domain"

PN MO200212341-A2.

PD 14-FEB-2002.

PD 03-AUG-2001; 2001WO-US24283.

PD 03-AUG-2000; 2000US-0632507.

PR (CORI-) CORIXA CORP.

PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Cheever MA, Gheysen D;

XX WPI: 2002-241743/29.

XX Her-2/neu fusion protein for treating or preventing cancer by eliciting

XX or enhancing an immune response to the protein, has Her-2/neu

XX extracellular domain fused to Her-2/neu intracellular or

XX phosphorylation domain

XX Claim 2; Fig 12; 14pp; English.

PS The present sequence is that of a fusion protein between the
 CC extracellular domain and phosphorylation domain of human Her-2/neu
 CC (see AAM51143), an oncogenic self-protein and target for anti-cancer
 CC vaccines. The fusion protein can be obtained by recombinant DNA
 CC methods. Her-2/neu overexpression correlates with a poor prognosis
 CC in breast and ovarian cancers. The invention provides Her-2/neu
 CC fusion proteins, nucleic acids encoding them, viral vectors, and
 CC vaccines comprising the fusion proteins or nucleic acid molecules.
 CC In preferred fusion proteins, the extracellular domain of a
 CC Her-2/neu protein is fused to a Her-2/neu intracellular domain or
 CC phosphorylation domain (or its DeltaD fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal ex vivo with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated
 CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.

XX Sequence 919 AA;

Query Match 68.8%; Score 2573.5; DB 21; Length 919;

PN	MOZ00020027-A2.
XX	
PD	13-APR-2000.
XX	
PF	05-OCT-1999; 99WO-DK00525.
XX	
PR	05-OCT-1998; 98DK-0001261.
PR	20-OCT-1998; 98US-0105011.
XX	
PA	(MEBT-) M & E BIOTECH AS.
XX	
PI	Steinaa L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Datum I;
PI	Gautam A, Birk P, Karlsson G;
XX	
DR	WPI: 2000-346917/30.
DR	N-PSTDB: AAA09455.
XX	
PT	Inducing immune responses to weakly immunogenic, tumor associated
PT	peptide antigens for the treatment of breast and prostate cancer
XX	
PS	Claim 62; Page 193-198; 220pp; English.
XX	
XX	This is the human heregulin 2 (Her2) sequence. Immunogenic analogues of
CC	Her2 can be used in the claimed method as an autovaccine to induce a CTL
CC	response. Subdominant CTL epitopes, antibody binding regions and
CC	cysteine residues involved in disulfide bonds are preserved in the
CC	immunogenized forms. Regions suitable for the insertion of foreign T
CC	helper epitopes were identified (see features table). The method
CC	is used for inducing immune responses against weakly immunogenic
CC	cell-associated peptide antigens (PA) such as those associated with
CC	cancers (self-proteins), e.g. human prostate specific membrane antigen
CC	(PSM), heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b).
CC	The method comprises effecting simultaneous presentation by antigen
CC	producing cells (APCs) of the animals immune system of: (1) at least 1
CC	CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC	B-cell group derived from the cell-associated PA; and (2) at least 1
CC	first T helper cell group which is foreign to the animal. Analogues of
CC	human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC	part of all known and predicted CTL and B-cell epitopes of the respective
CC	PA and including at least one foreign T helper epitope are also claimed.
CC	The method is used to treat prostate, prostate/breast or breast cancer
CC	when the PA is human PSM, FGF8b and Her2, respectively.
XX	
S0	Sequence 1255 AA;
	Query Match 64.5%; Score 2410.5; DB 21; Length 1255;
	Best Local Similarity 41.2%; Pred. No. 1.7e-142;
	Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1.
OY	35 LARGAASQVCGTGTDMKRLPASPETHLMDLHLVGOCGVGNGNELTLPTNASTSLFQ 94
Db	16 LPPGAASQVCGTGTDMKRLRPASPEHLDMLHLIVOGCGVGNLELTLLPTNASTSLFQ 75
OY	95 DIEOVGVYLIAHNOVROVPLORLRIVRGTQLPEDNYALAVLDNGDPLNTTPVTGASFG 154
Db	76 DIEOVGVYLIHNOVROVPLORLRIRVRTQLFEEDNYALAVALDNGDPNLNTPVTGASFG 135
OY	155 GLREVLNSLFTILGGVYLIOBNPOLCYODTILMKDIFFKNNOGLATLLDTNRSRACHFC 214
Db	136 GLREVLNSLFTILGGVYLIOBNPOLCYODTILMKDIFIKNNOLATLLIDTNRSRACHFC 195
OY	215 SPWGSGRCWGESSESDCOSLTRTVAGAGCARCKGPLPTDCHEQCACAGCTGRKHSDCLAC 274
Db	196 SPWGSGRCWGESSESDCOSLTRTVAGAGCARCKGPLPTDCHEQCACAGCTGRKHSDCLAC 255
OY	275 LHFNSGICELCPALYYNTDFESMRNPBERRYTFGASCYACPRNYLSTDVGS----- 329
Db	256 LHFNSGICELCPALYYNTDFESMPNBERRYTFGASCYACPYNYLSTDVGSCTIWC 315
OY	330 ----- 329
Db	316 PLIHNEVTAEDGTORCEKSCPCARVCYGLGMHLREVNAVYSANIQEFAGCKKIIFGSLA 375

[illegible]

Key	Location/Qualifiers
Region	774...782
/note="Antigenic epitope"	
WO200168677-A2.	
20-SEP-2001.	
16-MAR-2001; 2001WO-US40328.	
16-MAR-2000; 2000US-0527487.	
(GENZ) GENZYME CORP.	
Nicolette CA;	
WPI: 2001-616284/71.	
N-PSDB; AAD19731.	
Novel synthetic therapeutic compound for inducing immune response and for use in adoptive immunotherapy, has enhanced binding to major histocompatibility molecules and enhanced immunoregulatory properties	
Claim 4; Page 63-67; 69pp; English.	
The invention relates to synthetic therapeutic compounds (antigenic peptides) with enhanced binding to major histocompatibility complex (MHC) molecules and enhanced immunoregulatory properties relative to their natural counterparts. Compounds of the invention are useful for inducing an immune response in a subject and for use in adoptive immunotherapy. They are useful as components of anti-cancer vaccines and to expand immune effector cells that are specific for cancers characterized by expression of the breast cancer antigen, HER-2. Polynucleotides that encode peptides of the invention are useful as hybridisation probes and as primers for the detection of genes of gene transcripts that are expressed in antigen presenting cells (APCs), to confirm transduction of polynucleotides into host cells. The present sequence is human tyrosine kinase-type receptor, HER-2. compounds of the invention are designed based on the HER-2 antigenic peptide (774-782).	
Sequence 1255 AA;	
Query Match	64.5%; Score 2410.5; DB 22; Length 1255;
Best Local Similarity	41.2%; Pred. No. 1.7e-142;
Matches 510; Conservative 0; Mismatches 2; Indels 721; Gaps 1;	
35 LARGAASVGVCTGTDMKLRTPASPTHLDMLRHLVYGVCGVGNELTYLPTNASTSLFQ 94	
16 LPPGAASVGVCTGTDMKLRTPASPTHLDMLRHLVYGVCGVGNELTYLPTNASTSLFQ 75	
95 DIOEVGVYLIHNOVROYPLQRLRYRGTOLEFEDNYALAVLDNGDPLNNTTPYTGASPG 154	
76 DIOEVGVYLIHNOVROYPLQRLRYRGTOLEFEDNYALAVLDNGDPLNNTTPYTGASPG 135	
155 GIREVLQRLSTETILKGVYLIQRPOLCYODTILMKDIFPKNNQALATLIDTNRSRACHPC 214	
136 GIREVLQRLSTETILKGVYLIQRPOLCYODTILMKDIFPKNNQALATLIDTNRSRACHPC 195	
215 SPVCKGSRGWESSEDCOSLRTFVACGACRCKGPLPTDCCHBOCAAGCTGPKHSDDLAC 274	
196 SPVCKGSRGWESSEDCOSLRTFVACGACRCKGPLPTDCCHBOCAAGCTGPKHSDDLAC 255	
275 LAFNHSGLCELCPLALYNTDTFESMPNPERRYTFGASCVACRPNYLTSTVGS----- 329	
236 LAFNHSGLCELCPLALYNTDTFESMPNPERRYTFGASCVACRPNYLTSTVGSCTLVC 315	

OY	330	-----	329
Db	316	PLHNEVTAEDGTORCEKCSKPCARNCYGLGMEHLAEVNAVTSANTIOERAGCKKILFGSLA	375
OY	330	-----	329
Db	376	FLPESFDGDPASNTAPLQPEQLQVFTLEETGTLYISAMPDLPDLVFNQIAQVIRGR	435
OY	330	-----	329
Db	436	LHNGAVSLTLQGLGISWLGRLSLRELGSLALIHNTHLCEVHTVPMDQLFRNPQALLH	495
OY	330	-----	329
Db	556	NARHCLPCHPECOPNGSVTCFGPEADQVACAHYKDPPEVCYAKCPSPVLPDLSTYPIWK	615
OY	330	-----	329
Db	616	FPDEGACQPCPICTHSCVDLDDKGCRAQRASPLTSIYSANVGLLVVLGVYFGILL	675
OY	330	-----	329
Db	676	KRRQOKIRKTYMRRLQETELVEPLTPSGAMPNQAMRIKLELKKVKGSGAFGTYY	735
OY	330	-----	329
Db	736	KGIMIPDGENVKIPVAIKVLENTSPKANKELIDEATVMAVGSPVSVNLGICLTSTVQ	795
OY	330	-----	329
Db	796	LVTQLMPYGCLLDHVRENRRGLASODLNNCMQIAKMSYLEBDVRLVHRDLAARNVLKVS	855
OY	330	-----	329
Db	856	PNHWKTFDFGLARLLIDETEHADGKVPKYMMALESILLRRFTHQSDVMSYGVTWEL	915
OY	330	-----	329
Db	916	MTFGAKPYDGIPIAREIPDLLEKEGRBLPPICTIDVYIMVMKCMWIDSECRPRFRELVE	975
OY	330	-----	329
Db	976	FSRRARDPQRFVYIQNEDLGPASPLDSTFYRSLLEDDEKGLVDAAEYLVPOQGFECDP	1035
OY	330	--GAGGWNHNRHRSSTRSGGDLTLGLEPSEEEARSPRLAPSEGAQSDVFGDGLGMAA	387
Db	1036	APGAGWNHNRHRSSTRSGGDLTLGLEPSEEEARSPRLAPSEGAQSDVFGDGLGMAA	1095
OY	388	KGLOSLPTHDPSLQARYSEDPVYVLPBSETDGYVAPLTCSPQREYVQNQDVPKRPSPREG	447
Db	1096	KGLOSLPTHDPSLQARYSEDPVYVLPBSETDGYVAPLTCSPQREYVQNQDVPKRPSPREG	1155
OY	448	PLPARPAGATLERAKTILSPGKNGVYVDVAFGAVENPEYLLTPGGAPQHPPPAPSP	507
Db	1156	PLPARPAGATLERAKTILSPGKNGVYVDVAFGAVENPEYLLTPGGAPQHPPPAPSP	1215
OY	508	AFDNLVYWDODPPERGAPSPSTFGKPTAENPEYLGIDVP	546
Db	1216	AFDNLVYWDODPPERGAPSPSTFGKPTAENPEYLGIDVP	1254
RESULT 9			
AAB60167			
ID	AAB60167	standard; Protein: 1255 AA.	
XX	AAB60167:		
XX			
XX	03-APR-2001 (first entry)		
XX			

DE HER2 transgene plasmid construct encoded protein.
 XX Human: HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
 KW antibody.
 XX Homo sapiens.
 OS Synthetic.
 XX MO200100244-A2.
 XX 04-JAN-2001.
 XX 23-JUN-2000; 2000MO-US17229.
 XX 25-JUN-1999; 99US-0141316.
 PR 16-MAR-2000; 2000US-0189844.
 XX (GETH) GENENTECH INC.
 XX Erickson S, Schwall R;
 PI WPI: 2001-061962/07.
 DR N-PSDB; AAF24297.
 XX Treating tumors, particularly breast cancers, which overexpress an ErbB
 PT receptor and does not respond to an anti-ErbB antibody, comprises
 PT conjugating the antibody to a maytansinoid -
 XX
 XX Example 3; Fig 4; 92pp; English.
 XX The present invention provides a method of treating cancer by
 CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
 CC particular, the antibody is directed against ErbB2 (also known as HER2
 CC and p185neu). The method is particularly useful in the treatment of
 CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
 CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.
 XX
 XX Sequence 1255 AA:
 SO
 Query Match 64.5%; Score 2410.5; DB 22; Length 1255;
 Best Local Similarity 41.2%; Pred. No. 1,7e-142;
 Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

Db 436 LHNQAVSLTLOGISWLGSLRLSRELASGLALIHNNTHLCEVHTVPMQDLFRNPHQALLH 495
 Qy 330 ----- 329
 Db 496 TANRDECEVGEGLACHOLCARGHCWGPPTQVCNCSQPLRGQCEVECRVLQGLPREYV 555
 Qy 330 ----- 329
 Db 556 NARHCLPCHPECOQPNQSVTCGPEADQVACAHYKDPFCVARGCPGSKPDLSPYIMWK 615
 Qy 330 ----- 329
 Db 616 PPDEGACQPCPINCTHSCVDLDKCPAEQBRASPLTSTVSAVVGILLVVLGVVFGILI 675
 Qy 330 ----- 329
 Db 676 KRROQKIRKRYTMRLLQETELVEPLTPSGAMPNQAOMRLKETELRKVKVLGSGAFGTVY 735
 Qy 330 ----- 329
 Db 736 KGIWIPDGENKIPVAIVLRENTSPRANKELDEAVYMAVGSPYVRLGICLTSTVQ 795
 Qy 330 ----- 329
 Db 796 LVTQLMPYGLLDHVENRGRIGSODLLNMCQIAKGSYLEDVRLVHRDLAARNVLKS 855
 Qy 330 ----- 329
 Db 856 PNHVKITDFGLARLIDIDETEHADGKRPVKMMALESILRRRTHQSDVMSYGVTVWEL 915
 Qy 330 ----- 329
 Db 916 MTFGAKPYDGIIPAREIPDLLEKGERLPQPPICITIDVYIMVYKCMKIDSECRPRELYSE 975
 Qy 330 ----- 329
 Db 976 FSRMARDPQRFVYIQNEDLGPASPUDSTFYRSLIEDDMQDLVDAAEYLYPQCGFCFDP 1035
 Qy 330 ----- 329
 Db 1036 APGAGMWHHRSSSTSGGGDLTLGLEPSEEARPSPLAPSEAGSDVDEGDMKMAA 1095
 Qy 330 ----- 329
 Db 1096 KGIQSLPTHDSPLQRYSEDPVPLPSETDGVAVARLTSQPEYVNOQDVNRQPPSPREG 1155
 Qy 448 PLPAARPAATLERAKTSLSPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPAPASP 507
 Db 1156 PLPAARPAATLERAKTSLSPGKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPAPASP 1215
 Qy 508 AFDNLVYWDODPPERGAPSPFTKGPPTAENPEYLGLDVP 546
 Db 1216 AFDNLVYWDODPPERGAPSPFTKGPPTAENPEYLGLDVP 1254
 RESULT 10
 AAU74545
 ID AAU74545 standard; Protein: 1255 AA.
 XX AAU74545;
 AC 23-APR-2002 (first entry)
 DT
 XX
 DE Human HER2 (ErbB2) polypeptide.
 XX
 XX Human: HER2; ErbB; epidermal growth factor receptor; receptor;
 KW anti-ErbB antibody-maytansinoid conjugate; cancer; tumour; breast; ovary;
 KW stomach; endometrium; salivary gland; lung; kidney; colon; colorectum;
 KW thyroid; pancreas; prostate; bladder; ErbB2; neuronal disorder;
 KW glial disorder; astrocytal disorder; hypothalamic disorder;
 KW glandular disorder; macropneagal disorder; epithelial disorder;
 KW stromal disorder; blastocoele disorder; inflammatory disorder;
 KW angiogenic disorder; immunological disorder;

XX Homo sapiens.
 OS US2002001587-A1.
 PN
 XX
 PD 03-JAN-2002.
 XX
 PF 16-MAR-2001: 2001US-0811123.
 XX
 PR 16-MAR-2000: 2000US-189844P.
 PR 05-OCT-2000: 2000US-238327P.
 XX
 PA (ERIC/) ERICKSON S.
 PA (SCHM/) SCHWALL R.
 PA (SLIM/) SLIMKOWSKI M.
 XX
 PI Erickson S, Schwall R, Slimkowski M;
 XX
 DR WPI: 2002-163686/21.
 DR N-PSDB: ABR14058.
 XX
 XX
 PT Treating tumour characterised by overexpression of epidermal growth
 PT factor receptor, ErbB or cancer in mammal, comprises administering
 PT anti-ErbB antibody-maytansinoid conjugate to the mammal
 XX
 PS Example 3; Fig 7; 93pp; English.
 XX
 CC The invention relates to treating a tumour in a mammal, where the tumour
 CC is characterised by the overexpression of an epidermal growth factor
 CC receptor (ErbB) and does not respond or responds poorly, to treatment
 CC with an anti-ErbB antibody, comprising administering to the mammal an
 CC anti-ErbB antibody-maytansinoid conjugate. The method is useful for
 CC treating cancer or tumours of the breast, ovary, stomach, endometrium,
 CC salivary gland, lung, kidney, colon, colorectum, thyroid, pancreas,
 CC prostate and bladder, preferably breast cancer. The breast cancer is a
 CC metastatic breast cancer or an aggressive form of metastatic breast
 CC cancer which overexpresses ErbB2. The method is also useful for treating
 CC neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,
 CC epithelial, stromal, blastocellic, inflammatory, angelioma and
 CC immunological disorders. This sequence represents the human HER2 (ErbB2)
 CC polypeptide of the invention.
 CC
 SO Sequence 1255 AA:
 XX
 Query Match 64.5%; Score 2410.5; DB 23; Length 1255;
 Best Local Similarity 41.2%; Pred. No. 1.7e-142;
 Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

Db 376 FLPSFDGASNTAPLQPOLOVETLLEITGYLISAMPDSLPLSVFQNLQVINGRI 435
 QY 330 ----- 329
 Db 436 LHNCAVSLTLQGLGISWGLRLSELGSLALIHNTHLCEVHTVPWDQLEFRPHQALLH 495
 QY 330 ----- 329
 Db 496 TANRPEDECVGEGLACHQCLARGHCMPPTQVCNCSQFLRGCEVECEVLOGLPREYV 555
 QY 330 ----- 329
 Db 556 NARHCLPHECOPQNSVTCFGEADQVCAHAHYKDPFCVARGPSGVKPDLSYPMWK 615
 QY 330 ----- 329
 Db 616 FPDEGACQPPINCTHSCVDLDDKGPARGASPLTSYSAVVGILLVYLVGVPBILI 675
 QY 330 ----- 329
 Db 676 KRRQAKIRKYTMRLLOETELVEPLTPSGAMPNQAMRLKETELRKVKVSGAGFTVY 735
 QY 330 ----- 329
 Db 736 KGIWIPDENVKIPVAIKVLRNTSPKANKELDEAVVMAGVSPYVSRLLGICLTSTVQ 795
 QY 330 ----- 329
 Db 796 LVYQMLPYGCLLDHVRNRRRLGSLQDLNMCQIAKMSYLEDRVLRHDLAARNLVKS 855
 QY 330 ----- 329
 Db 856 PNHYKITDFGLARLDIDETEHADGKVPDKMALESLRRRPTHOSDYVSYGVYMWEL 915
 QY 330 ----- 329
 Db 916 MTEGAKPYDGIPIAREIPDLLEKGERLPDPICTIDYIMVWKMDSECPRELYSE 975
 QY 330 ----- 329
 Db 976 FSRWARPQRFVYIQNEDLPASPLDSTFYRSLEDDMDGLVDAEYLVPOQGFPCPD 1035
 QY 330 --GAGGVNHHRRNSSYRSGGDLTLCLEPSEEPAPSPAPSGACSDYVDGLGAA 387
 Db 1036 APGAGGVNHHRRNSSYRSGGDLTLCLEPSEEPAPSPAPSGACSDYVDGLGAA 1095
 QY 388 KGLQSLPTHPSPLQRYSEDPTVPLPSETDGYVAPLCSFQPEYVNOPDVROPSPREG 447
 Db 1096 KGLQSLPTHPSPLQRYSEDPTVPLPSETDGYVAPLCSFQPEYVNOPDVROPSPREG 1155
 QY 448 PLPARAPAGATLERAKTLPSPKNGVVDVAFGAVENPEYLPQGGAPQHPAPFSP 507
 Db 1156 PLPARAPAGATLERAKTLPSPKNGVVDVAFGAVENPEYLPQGGAPQHPAPFSP 1215
 QY 508 AFNLLYTWDDPPEBGRAPSTFKGTPTAENPEYLGIDVP 546
 Db 1216 AFNLLYTWDDPPEBGRAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 11
 AAW01111
 ID AAW01111 standard; Protein: 1255 AA.
 XX
 AC AAW01111:
 XX
 DT 01-JAN-1997 (first entry)
 XX
 DE HER-2/neu protein.
 XX
 KW HER-2/neu; c-erbB1; p185; oncogene; tyrosine protein kinase;
 KW breast cancer; ovary cancer; colon cancer; lung cancer;
 KW prostate cancer; immunisation; tumour; vaccine; vector.

XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FH 676..1255
 FT Domain /label=intracellular_domain
 FT /note="claimed domain, useful for immunisation"
 XX
 PN M09630514-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 28-MAR-1996; 96WO-US01689.
 XX
 PR 31-MAR-1995; 95US-0414417.
 XX
 PA (UNIW) UNIV WASHINGTON.
 PI Cheever MA, Disls ML;
 XX
 DR WPI: 1996-455361/45.
 DR N-PSDB; AAT40739.
 XX
 PT DNA encoding HER-2/neu poly:peptide(s) - used for prevention or
 PT treatment of malignancies with which the HER-2/neu oncogene is
 PT associated
 XX
 PS Claim 2; Page 56-61; 71pp; English.
 XX
 CC Human HER-2/neu protein (AAM01111), also called p185 or c-erbB2, is
 CC the product of the HER-2/neu oncogene (see also AAT40739). The
 CC protein is over-expressed in various cancers, including breast,
 CC ovarian, colon, lung and prostate. The intracellular domain of the
 CC protein can be used to immunise an animal against a malignancy with
 CC which the oncogene is associated. The polypeptide can be produced
 CC in transformed host cells for use in immunisation. Alternatively,
 CC animal cells are transfected in vivo or ex vivo with a viral vector
 CC that directs expression of the polypeptide.
 CC
 XX
 SO Sequence 1255 AA;
 Query Match 64.3%; Score 2405.5; DB 17; Length 1255;
 Best Local Similarity 41.1%; Pred. No. 3.5e-142;
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;
 QY 35 LARGAASOVCTGTMTKRLRASPETHLDMKRLHLYQGCGVQVGNLELYLPNASTSLFLQ 94
 DB 16 LPPGAASOVCTGTMTKRLRASPETHLDMKRLHLYQGCGVQVGNLELYLPNASTSLFLQ 75
 QY 95 DIOEVOGVLLAHNOVROVPLQRLRIVRGTOLEEDNYALAVLNDGDPNNTPTVGASPG 154
 DB 76 DIOEVOGVLLAHNOVROVPLQRLRIVRGTOLEEDNYALAVLNDGDPNNTPTVGASPG 135
 QY 155 GLRELQLRSLTEILKGLVLIQRNPOLCYODTILMKDIFHKNNQALTLIDTNRSRACHPC 214
 DB 136 GLRELQLRSLTEILKGLVLIQRNPOLCYODTILMKDIFHKNNQALTLIDTNRSRACHPC 195
 QY 215 SPMCKGSGMCKWSSBDCSLRTTVACAGCARCKGPLPTDCCHEGCAAGCTGKRHSDCLAC 274
 DB 196 SPMCKGSGMCKWSSBDCSLRTTVACAGCARCKGPLPTDCCHEGCAAGCTGKRHSDCLAC 255
 QY 275 LAFNHSIGICELCPALVYNTDTFESMPNPEGRYTFGASCVTACVYNTLSTVGS----- 329
 DB 256 LAFNHSIGICELCPALVYNTDTFESMPNPEGRYTFGASCVTACVYNTLSTVGSCTIVC 315
 QY 330 ----- 329
 DB 316 PLHNOEVAEDGTQRCCKSPCARVYCYGLMEHLREAVATSANIOEFAGCKITFGSLA 375
 QY 330 ----- 329
 DB 376 FLPEFDDPASPNTAPLOPELOVFEETLEITGYLISAMPDLPDLVSFONLQVIRGRI 435

QY 330 ----- 329
 DB 436 LHNCAVSLTLQCGISWIGLSRLBSGLALIHNNHLCFVHTVPMQDLFRNPQALLH 495
 QY 330 ----- 329
 DB 496 TANREDECBVGBGLACHQLCARGHCWGPPTQCVNCSQFLRQCECBECRVLGLPREV 555
 QY 330 ----- 329
 DB 556 NARHCLPCHPEBCQPNQSVTCFGEADOCVACAHYKDPFCVAPCPSGVKPDLSYPIWK 615
 QY 330 ----- 329
 DB 616 PPDEGACQPCPINCTHSCVDLDDKGCRAEQNASPLTISAVGILLVVLGVVFGILI 675
 QY 330 ----- 329
 DB 676 KRQCKIRKRYTMRLLQETELVEPLTPSGAMPNQAOMRLKETELRKYKVLASGAFGVY 735
 QY 330 ----- 329
 DB 736 KGIWIPDEENKIPVAIKVIRENTSPKANKETLDEAVYVAGVSPYSRLIGICTSTVQ 795
 QY 330 ----- 329
 DB 796 LVTOAMPYGCCLLDHYRENKRLGSODLNMCKQIAKMSYLEDVRLVHRDLAARVLYKS 855
 QY 330 ----- 329
 DB 856 PNHVKTDFGLARLIDIDETEHADGKVPKIMMALESILRRRFTHQSDVMSGYTVNEL 915
 QY 330 ----- 329
 DB 916 MTFGAKPYDGIIPAREIPDLLEKGERLPORPCTIDVYIMVCKMIDSECRFRRELVE 975
 QY 330 ----- 329
 DB 976 FSRMARDPQRFVYIQNEDLGRASPLDSTFFYRSLEDDMDGLVDABEYLVPOGFFCPR 1035
 QY 330 --GAGGVNHHRRSSSTSSGGDLTLGLPSESEAPRSLPABSGASDVDFDGLMGAA 387
 DB 1036 APGAGGMVHHRSSSTSSGGDLTLGLPSESEAPRSLPABSGASDVDFDGLMGAA 1095
 QY 388 KGLOSLPTHPDPSLQRYSEDDPYPLPSETDGVVAPLTCSPQEVYVNOPDVPKPPSPREG 447
 DB 1096 KGLOSLPTHPDPSLQRYSEDDPYPLPSETDGVVAPLTCSPQEVYVNOPDVPKPPSPREG 1155
 QY 448 PLPARPAGATLERAKTLPCKNGVYKDFAFGAVENPEVLTPOGGAAPQPHPPAPSP 507
 DB 1156 PLPARPAGATLERAKTLPCKNGVYKDFAFGAVENPEVLTPOGGAAPQPHPPAPSP 1215
 QY 508 AFDNLYWDDPPERGAPPSFTFGPTAENPEYLGLDVP 546
 DB 1216 AFDNLYWDDPPERGAPPSFTFGPTAENPEYLGLDVP 1254
 RESULT 12
 AAM92406
 ID AAM92406 standard; Protein: 1255 AA.
 XX
 XX AAM92406;
 AC
 AC
 AC
 DT 21-Apr-1999 (first entry)
 XX
 DE Human HER-2/neu oncogene protein.
 XX
 XX HER-2/neu; oncogene; immune response; T cell; B cell; immunisation;
 KW malignancy; treatment; tumour.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH

FT Region 676..1255
 /note= "Region which elicits immune response"
 XX US5869445-A.
 PN 09-FEB-1999.
 PD 01-APR-1996; 9605-0625101.
 PF 01-APR-1996; 9605-0625101.
 PR 01-APR-1996; 9605-0625101.
 PR 17-MAR-1993; 9305-0033644.
 PR 12-AUG-1993; 9305-0106112.
 PR 31-MAR-1995; 9505-0414417.
 XX (UNIM) UNIV WASHINGTON.
 PA Cheever MA, Dists ML;
 PI
 XX
 XX WPI; 1999-152835/13.
 DR N-PSDB; AAX01912.
 XX
 XX Use of HER-2/neu polypeptides - for eliciting an immune response to
 PT an HER-2/neu associated malignancy, particularly for treating or
 PT preventing tumours
 XX
 XX Claim 3; Column 31-38; 26pp; English.
 PS
 CC This sequence represents the human HER-2/neu oncogene protein. A fragment
 CC of this protein is used in a method for eliciting or enhancing an immune
 CC response to HER-2/neu protein. The polypeptide can stimulate T cells and
 CC B cells to produce an immune response to the HER-2/neu protein. The
 CC method can be used for immunisation against a malignancy in which the
 CC HER-2/neu oncogene is associated and in the treatment of an existing
 CC tumour, or to prevent tumour occurrence or reoccurrence.
 CC
 XX Sequence 1255 AA:
 SO
 Query Match 64.38; Score 2405.5; DB 20; Length 1255;
 Best Local Similarity 41.18; Pred. No. 3.5e-142;
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 330 ----- 329
 Db 496 TANPEDECVGEGLACHQLCARGHCWGPPTQVCNCSQFLRGQCEVCECRYLQGLPREYV 555
 QY 330 ----- 329
 Db 556 NARHCLEPCHECQFQNSVYCFGEADQCAAHYKDPPECVARGSPGVKPDLSYMPWK 615
 QY 330 ----- 329
 Db 616 FPDEGACQPCPINCTHSCVDLDDKGCAPQARASPLTISAVGILLVVLGVVFGIL 675
 QY 330 ----- 329
 Db 676 KRRQOKIRKYTMRLLOETELVEPLPSGAMPNQAQHRILKETELRRKVKVLSGAGFTVY 735
 QY 330 ----- 329
 Db 736 KGIVPDGENVKIIVAIKVLRENTSPANKETLDEAVYMGVGSPIYSRLIGICTSTVQ 795
 QY 330 ----- 329
 Db 796 LVYQLMPYGCLLDHVRENRRGLSQDLLMCMQJAKGMSYLEDVRLVHRDLARNVLVKS 855
 QY 330 ----- 329
 Db 856 PNHYKITDFGLARLLDIDETEHADGKVPKIMNALESILRRPETHOSDVNSYGVTWEL 915
 QY 330 ----- 329
 Db 916 MTFCAKPYDGIPIAREIPLDLEKGRLPQPICTIDVYIMVYKCMIDSECRPRELVE 975
 QY 330 ----- 329
 Db 976 FSRMARDPQRFVYIQNEDLGPASPLDSTFYRSLLDEDDMGDLVDAEELVYPOGFPCDP 1035
 QY 330 ----- 387
 Db 1036 APGAGGVNHHRRSSSTRSGGDLTGLPESEEARPSPLASGASDVFDGGLGMAA 1095
 QY 388 KGLSLPTHPSPQRYSEDPVPLPSETDGYVAPLTCSPQREYVNOPDVAPPPSPREG 447
 Db 1096 KGLSLPTHPSPQRYSEDPVPLPSETDGYVAPLTCSPQREYVNOPDVAPPPSPREG 1155
 QY 448 PLPAARAGATLERAKTLPCKNGVVKDVFAGAVENPEYLTPQGAAPQHPPPAFSP 507
 Db 1156 PLPAARAGATLERAKTLPCKNGVVKDVFAGAVENPEYLTPQGAAPQHPPPAFSP 1215
 QY 508 AFQNLVYWDQDPERGAPSPSTFKGTPAENPEYLGIDVP 546
 Db 1216 AFQNLVYWDQDPERGAPSPSTFKGTPAENPEYLGIDVP 1254

RESULT 13
 ID AAB21198 standard; protein; 1255 AA.
 XX AAB21198;
 AC AAB21198;
 XX
 DT 12-JAN-2001 (first entry)
 XX
 DE Human HER-2/neu protein.
 XX
 KW Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;
 KW breast cancer; prostate cancer; ovarian cancer; lung cancer;
 KW colon cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200044899-A1.
 XX
 PD 03-AUG-2000.
 XX

PF 28-JAN-2000; 2000MO-US02164.
 XX
 PR 29-JAN-1999; 99US-0117976.
 XX
 PA (CORI-) CORIXA CORP.
 PA (SMIK) SMITHKLINE BEECHAM.
 XX
 PI Cheever MA, Gheysen D;
 XX
 DR WPI; 2000-505976/45.
 DR N-PSDB; AAA89736.
 XX
 PT HER-2/neu extracellular domain/phosphorylation domain fusion proteins
 PT useful for vaccinating against breast, ovarian, colon, lung and
 PT prostate cancers -
 PS
 PS Claim 52; Fig 7; 128pp; English.
 XX
 CC The present sequence is the human HER-2/neu protein. It is a member of
 CC the tyrosine kinase family of receptor-like glycoproteins and shows
 CC homology to the epidermal growth factor receptor (EGFR). It probably
 CC plays a part in cell growth and/or differentiation. The HER-2/neu
 CC gene is an oncogene. An HER-2/neu fusion protein comprising a
 CC HER-2/neu extracellular domain fused to a HER-2/neu phosphorylation
 CC domain may be used to treat or prevent cancer by eliciting or
 CC enhancing an immune response to the HER-2/neu protein. It may be used
 CC to treat malignancies such as breast, ovarian, colon, lung and
 CC prostate cancers, and may be used as an antigen to vaccinate against
 CC these neoplasias.
 CC
 XX
 SQ Sequence 1255 AA;
 Query Match 64.3%; Score 2405.5; DB 21; Length 1255;
 Best Local Similarity 41.1%; Pred. No. 3.5e-142;
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

Db 556 NARHCLPCHPECQPNQSVTCGPEADQVACAHYKIDPPFCVAPRPSGVKPDLSYMPIWK 615
 QY 330 ----- 329
 Db 616 FPDEGACQPCRPINCTHSCVDLDKGPAPORASPLTISIAYVIGLLVVLGVVFGILI 675
 QY 330 ----- 329
 Db 676 KRROOKIKRYTMRLLQETELVEPLTPSGAMPNOAQMRLKETELRKVKYLGSQAFGVY 735
 QY 330 ----- 329
 Db 736 KGIWIDGENVKIPVAIKVIRENTSPKANKETLDEAYVMAGVSPYVSRLLGICLTSTVQ 795
 QY 330 ----- 329
 Db 796 LVTLMPYCLLDHYRENHRLGSDLLNMCQIAKMSYLEDVRLVHRDLAARVLYKS 855
 QY 330 ----- 329
 Db 856 PNHVKITDFGLARLLDIDETEHADGCKVPIKMALESILRRRTHQSDVMSYCVTVNEL 915
 QY 330 ----- 329
 Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPICTIDVYIMVNCWMIDSECRPRFRELVS 975
 QY 330 ----- 329
 Db 976 FSRMARDPQRFVYIIONEDLGAPSLDSTFYRSLLLEDDMGDLVDAEYLVPOQGFCCPD 1035
 QY 330 --GAGWVNHHRSSSTRSGGDLTLGLEPSEEARSPPLAPSEAGSDVFDGDLGMCA 387
 Db 1036 APGAGWVNHHRSSSTRSGGDLTLGLEPSEEARSPPLAPSEAGSDVFDGDLGMCA 1095
 QY 388 KGLSLPTNDSPLQRYSEDPVPLPSETDGYVAPLTCSPQREYVNPQDVAPRPPREG 447
 Db 1096 KGLSLPTNDSPLQRYSEDPVPLPSETDGYVAPLTCSPQREYVNPQDVAPRPPREG 1155
 QY 448 PLPAARPGATLERAKTISPCKNGVYKDFAPFGAVENPEYLTPOGGAAPRHPARFSP 507
 Db 1156 PLPAARPGATLERAKTISPCKNGVYKDFAPFGAVENPEYLTPOGGAAPRHPARFSP 1215
 QY 508 AFDNLVYWDODPPERKAPSTFKGTPTAENPEYLGIDVP 546
 Db 1216 AFDNLVYWDODPPERKAPSTFKGTPTAENPEYLGIDVP 1254

RESULT 14
 AAY84780
 ID AAY84780 standard; Protein: 1255 AA.
 XX
 AC AAY84780;
 XX
 DT 08-AUG-2000 (first entry)
 XX
 DE Amino acid sequence of the SPLICE erbb-2 receptor protein.
 XX
 KW SPLICE erbb-2 receptor protein; cell transformation disorder; cancer;
 KW tumor cell proliferation; tissue degeneration; arthropathy;
 KW bone resorption; inflammatory disease; degenerative disorder;
 KW wound healing.
 XX
 OS Homo sapiens.
 XX
 PN WO200020579-A1.
 XX
 PD 13-Apr-2000.
 XX
 PE 01-Oct-1999; 99WO-CA00912.
 XX
 PR 02-Oct-1998; 98US-0165192.
 XX

XX New antigen-presenting cells, useful as vaccines for eliciting or
 PT enhancing an immune response to HER-2/neu protein, particularly useful
 PT for treating or preventing cancer, e.g. breast cancer
 PS Claim 2; Page 41-46; 49pp; English.
 XX The invention provides an isolated antigen-presenting cell, which
 CC expresses at least an immunogenic portion of a polypeptide that produces
 CC an immune response to HER-2/neu protein. The antigen-presenting cells are
 CC useful as vaccines for eliciting or enhancing an immune response to
 CC HER-2/neu protein, particularly in treating or preventing malignancies in
 CC which the HER-2/neu oncogene is associated. Specifically, these are
 CC useful for treating or preventing cancer, e.g. breast cancer, ovarian,
 CC colon, lung or prostate cancers. The present sequence represents
 CC the human HER-2/neu protein (also known as p185 or c-erbB2).
 XX
 SQ Sequence 1255 AA;
 Query Match 64.3%; Score 2405.5; DB 22; Length 1255;
 Best Local Similarity 41.1%; Pred. No. 3.5e-142;
 Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY 35 LARGAATQVCTGDMKRLPASPEHLMRLHLYGCGVQVGNIELTYLPTNASLFLQ 94
 DB 16 LPPGAATQVCTGDMKRLPASPEHLMRLHLYGCGVQVGNIELTYLPTNASLFLQ 75
 QY 95 DIOEVQVLLAHNOVROVRLRIYRGTLQFEDNVALLVNDNDPLNNTPTVGSAPG 154
 DB 76 DIOEVQVLLAHNOVROVRLRIYRGTLQFEDNVALLVNDNDPLNNTPTVGSAPG 135
 QY 155 GLRELQLSLTEILKGVLIQIRNPOLCYQDTILMKDIFHKNNQLATLIDTNRSHACHP 214
 DB 136 GLRELQLSLTEILKGVLIQIRNPOLCYQDTILMKDIFHKNNQLATLIDTNRSHACHP 195
 QY 215 SPMCKSGKMGESSEDCOSTRTVCAGGACARCKGRLPTDCCHQCAGCTGPKHSDCIAC 274
 DB 196 SPMCKSGKMGESSEDCOSTRTVCAGGACARCKGRLPTDCCHQCAGCTGPKHSDCIAC 255
 QY 275 LHFNSGICELHCPALVYVNDTFESMPNPGRYTFEGASCTACPYNTLSDVGS 329
 DB 256 LHFNSGICELHCPALVYVNDTFESMPNPGRYTFEGASCTACPYNTLSDVGS 315
 QY 330 329
 DB 316 PLHNGEVTAEDEGTQCEKSKPCARVCGLGMEHLREVRATVANSIOEFAGCKITFGSLA 375
 QY 330 329
 DB 376 FLPSFDGDPASNTAPLQPEQLQVEFTEITGYLYISAMPDLSPLSVQNLQVIRGRI 435
 QY 330 329
 DB 436 LHNQAVSLTLQGLISWLGLSLRELGLSGLALIHNNHLCFVHTVPMQDLFRNPQALLH 495
 QY 330 329
 DB 496 TANRPEDCEVGEGLACHOLCARGHCHMGPTQCVNCSQFLRGQCEVEECRYLQGLPREV 555
 QY 330 329
 DB 556 NARHCLPCHPCOPONGSVTCFGEADQVACAHYKDRPFCVARGPSGVKPDLSYMPIWK 615
 QY 330 329
 DB 616 FPDEGACQPCPINCTHSCVDLDKGPABQASPLTSIIASVVGILLVVLGVVFGILI 675
 QY 330 329
 DB 676 KRROOKIKKTYMRLLQETELVEPLTPSGANPQAOQRIKTELRKVKVVLGSGAFGIVY 735
 QY 330 329

DB 736 KGIWIPDGENVKIPVAIKVIRENTSPKANKIELDEAYVMAGVSBVYSRLGICLTSTVQ 795
 QY 330 329
 DB 796 LVTLQMPYGCILLDHVRENKRLGSODLLNMCQIAKMSYLEDVRLVHRDLAARNVLVKS 855
 QY 330 329
 DB 856 PNHYKITDFGLARLLIDIDETEVHADGKVPKMMALSLARRRPTHOSDVMSYGVTVWEL 915
 QY 330 329
 DB 916 MTFGAKPYDGIPIAREIDPLEKGERLPQPICTIDVYMINVKCMWIDSECRPRFRELIVSE 975
 QY 330 329
 DB 976 FSRMARDPQREVVIONEDLGASPLDSTFTFRSLLEDDMGDLVDAEYLVPOGFFCPDP 1035
 QY 330 387
 DB 1036 APGAGGVNHHNRSSSTRSGGDLTLGLEPSEEAAPRSPLAPSEAGSDVFDGLMGAA 1095
 QY 388 KGLQSLPTHDSPLOIRYSEDTVPPLPSETDGVAPLTCSPQREYVNOVDVAPRPPSPREG 447
 DB 1096 KGLQSLPTHDSPLOIRYSEDTVPPLPSETDGVAPLTCSPQREYVNOVDVAPRPPSPREG 1155
 QY 448 PLPAARPAATLERAKTLSPGKNGVVDVAFGAVENPEYLTPOGGAAPRPPPAFSP 507
 DB 1156 PLPAARPAATLERAKTLSPGKNGVVDVAFGAVENPEYLTPOGGAAPRPPPAFSP 1215
 QY 508 AFDNLVYWDODPPERKAPSTFKGTPTAENPEYLGLDVP 546
 DB 1216 AFDNLVYWDODPPERKAPSTFKGTPTAENPEYLGLDVP 1254

Search completed: April 28, 2003, 13:39:43
 Job time : 72.3846 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: April 28, 2003, 13:38:09 ; Search time 7.98328 Seconds

(Without alignments)
2613.108 Million cell updates/sec

Title: US-09-821-883-25

Sequence: 1 GAGGWHHRHRRSSSTRSGG.....STFKGPTAENPEYLGLDVP 217

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1177	99.6	1255	1 A24571	protein-tyrosine k
2	966.5	81.8	1254	1 A48161	p-185 precursor -
3	954	80.7	1260	1 TVRIND	protein-tyrosine k
4	164	13.9	1210	1 GQHUE	epidermal growth f
5	157	13.3	1308	2 A47253	epidermal growth f
6	145.5	12.3	707	2 A46302	PTB-associated spl
7	138	11.7	698	1 TVEFVY	protein-tyrosine k
8	138	11.7	1223	1 TYCHIV	epidermal growth f
9	137.5	11.6	392	1 PTHU66	salivary proline-r
10	135.5	11.5	553	1 H70786	probable dihydroli
11	135	11.4	620	2 S06733	hydroxyproline-ric
12	134.5	11.4	379	2 S31719	proline-rich prote
13	134	11.3	309	2 T29293	hypothetical prote
14	134	11.3	674	2 T05264	probable serine/th
15	134	11.3	1210	2 A53183	epidermal growth f
16	132.5	11.2	322	2 S25299	extensin precursor
17	132.5	11.2	442	2 S50062	cell wall glycopro
18	131.5	11.1	473	2 S50755	hypothetical prote
19	130	11.0	847	2 P96531	hypothetical prote
20	129.5	11.0	251	1 PTHU66	salivary proline-r
21	129.5	11.0	476	2 T27051	hypothetical prote
22	129	10.9	440	2 TC7807	Wiskott-Aldrich sy
23	127.5	10.8	403	2 S27296	PLP2 protein - hu
24	127.5	10.8	3149	1 OQBE8	BPFL1 protein - hu
25	127	10.7	369	2 S20500	hydroxyproline-ric
26	126	10.7	990	2 T14756	hypothetical prote
27	126	10.7	4957	2 T03455	AKR protein - huma
28	126	10.7	5262	2 T03454	AKR protein - huma
29	125.5	10.6	760	2 T06291	extensin homolog T

30	125	10.6	1460	1 EDBE1F	Immediate-early pr
31	124.5	10.5	775	1 EDBE11	Immediate-early pr
32	124	10.5	551	2 S57447	HPBR1-7 protein -
33	124	10.5	1106	1 TVHUGL	transforming prote
34	123.5	10.4	1494	2 T14355	protein-tyrosine-P
35	123	10.4	2715	2 T13049	eyella - fruit fly
36	122	10.3	656	2 A41870	dna protein - Str
37	121.5	10.3	300	2 J02220	hydroxyproline-ric
38	121.5	10.3	368	2 C29356	hydroxyproline-ric
39	121.5	10.3	938	1 Q0BE24	nuclear antigen EB
40	121.5	10.3	1357	2 T29265	hypothetical prote
41	121	10.2	1585	2 T31611	hypothetical prote
42	120.5	10.2	301	2 T21314	hypothetical prote
43	120.5	10.2	424	2 T33663	hypothetical prote
44	120.5	10.2	449	2 S16748	proline-rich prote
45	120	10.2	318	2 T29479	hypothetical prote

ALIGNMENTS

RESULT 1
A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; Kinase-related transforming protein e
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence, revision 06-Dec-1996 #text, change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Salto, T
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth.
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epid
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Coonsens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chro
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALM', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptio
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:q183983; PIDN:AAA58637.1; PID:9553332
 C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
 C:Genetics:
 A:Gene: GDB:ERB2; NGL: NEU; HER-2
 A:Cross-references: GDB:120613; OMIM:164870
 A:Map position: 17q21.1-17q21.1
 A:Introns: 25/1; 75/3; 147/1; 883/3
 A>Note: the list of introns is incomplete
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-125/Domain: signal sequence #status predicted <SIG>
 F:22-653/Domain: extracellular #status predicted <EXT>
 F:70-304/Domain: EGF receptor extracellular domain repeat <E1>
 F:395-605/Domain: EGF receptor extracellular domain repeat <E2>
 F:554-675/Domain: transmembrane #status predicted <TM>
 F:576-1235/Domain: intracellular #status predicted <INT>
 F:718-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 F:68,124,187,259,530,571,629/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:86/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:753/Active site: Lys #status predicted
 F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 99.6%; Score 1177; DB 1; Length 1255;
 Best Local Similarity 99.5%; Pred. No. 1,4e-69;
 Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAK 60
 DB 1038 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAK 1097

QY 61 LOSLPHPDPSPLQRYSDPPLPSEFDGVVAPLTCSPQEVYNQDPVRQPPSPRG 120
 DB 1098 LOSLPHPDPSPLQRYSDPPLPSEFDGVVAPLTCSPQEVYNQDPVRQPPSPRG 1157

QY 121 PAARPGATLERAKTSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAP 180
 DB 1158 PAARPGATLERAKTSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAP 1217

QY 181 DNLTYWDQDPERGAPSTFKGPTAENPEYLTGLDVP 217
 DB 1218 DNLTYWDQDPERGAPSTFKGPTAENPEYLTGLDVP 1254

RESULT 2
 I48161
 P-185 precursor - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Jun-1999
 C:Accession: I48161
 R:Nakamura, T.; Ushijima, T.; Ishizaka, Y.; Nagao, M.; Arai, M.; Yamazaki, Y.; Ishikawa, Gene 140, 251-255, 1994
 A:Title: Cloning and activation of the Syrian hamster neu proto-oncogene.
 A:Reference number: I48161; MUID:94193007; PMID:7908275
 A:Accession: I48161
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1254 <RES>
 A:Cross-references: GB:D16295; NID:9493236; PIDN:BA03801.1; PID:9747595
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP
 F:118-983/Domain: protein kinase homology <KIN>
 F:726-734/Region: protein kinase ATP-binding motif
 Query Match 81.8%; Score 966.5; DB 2; Length 1254;
 Best Local Similarity 82.0%; Pred. No. 8,1e-56;
 Matches 178; Conservative 13; Mismatches 25; Indels 1; Gaps 1;

QY 1 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAK 60
 DB 1038 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAK 1097

QY 61 LOSLPHPDPSPLQRYSDPPLPSEFDGVVAPLTCSPQEVYNQDPVRQPPSPRG 120
 DB 1098 LOSLPHPDPSPLQRYSDPPLPSEFDGVVAPLTCSPQEVYNQDPVRQPPSPRG 1157

QY 121 PAARPGATLERAKTSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAP 180
 DB 1158 PAARPGATLERAKTSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAP 1216

QY 181 DNLTYWDQDPERGAPSTFKGPTAENPEYLTGLDVP 217
 DB 1217 DNLTYWDQDPERGAPSTFKGPTAENPEYLTGLDVP 1253

RESULT 3
 TVRNU
 protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
 C:Accession: A24562; A61204
 R:Barthmann, C.I.; Hung, M.C.; Weinberg, R.A.
 Nature 319, 226-230, 1986
 A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein
 A:Reference number: A24562; MUID:86118662; PMID:3945311
 A:Accession: A24562
 A:Molecule type: mRNA
 A:Residues: 1-1260 <BAR>
 A:Cross-references: EMBL:X03362; NID:956745; PIDN:CAA27059.1; PID:956746
 R:Masu, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen Carcinogenesis 12, 1975-1978, 1991
 A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals a 2-thiazolyl]formamide or N-methyl-N-nitrosourea
 A:Reference number: A61204; MUID:92035293; PMID:1682063
 A:Accession: A61204
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 637-663, 'V', 665-702 <MAS>
 A>Note: authors translated the codon GCA for residue 25 as Val
 C:Genetics:
 A:Gene: neu
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phosphatase
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1260/Domain: signal sequence #status predicted <SIG>
 F:558-680/Domain: transmembrane #status predicted <TM>
 F:723-988/Domain: protein kinase homology <KIN>
 F:731-739/Region: protein kinase ATP-binding motif
 F:71,191,263,535,576,634/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:691/Binding site: phosphate (Thr) (covalent) #status predicted
 F:758/Active site: Lys #status predicted
 F:882,1227,1253/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 80.7%; Score 954; DB 1; Length 1260;
 Best Local Similarity 81.6%; Pred. No. 5,4e-55;
 Matches 177; Conservative 8; Mismatches 32; Indels 0; Gaps 0;

QY 1 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAK 60
 DB 1043 GAGCWNHRRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAK 1102

QY 61 LOSLPHPDPSPLQRYSDPPLPSEFDGVVAPLTCSPQEVYNQDPVRQPPSPRG 120
 DB 1103 LOSLPHPDPSPLQRYSDPPLPSEFDGVVAPLTCSPQEVYNQDPVRQPPSPRG 1162

QY 121 PAARPGATLERAKTSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAP 180
 DB 1163 PAARPGATLERAKTSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPAPSPAP 1222

QY 181 DNLTYWDQDPERGAPSTFKGPTAENPEYLTGLDVP 217

Db 1223 DNLTYMDNSSEOCPPSNFEGTPTAENPEYLGDVP 1259

RESULT 4

Epidermal growth factor receptor precursor - human

N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erDB

C:Species: Homo sapiens (man)

C:Date: 15-Nov-1984 #sequence_revision 27-Nov-1985 #text_change 11-Jun-1999

C:Accession: A00641; A25772; S30024; A38672; A00642; A43615; A23062; A05281; A60143; A317; P.H.

R:Ulrich, A.; Coussens, L.; Hayflick, J.S.; Dull, T.J.; Gray, A.; Tam, A.W.; Lee, J.; Y

Nature 309, 418-425, 1984

A:Title: Human epidermal growth factor receptor cDNA sequence and aberrant expression of

A:Accession number: A00641; MUID:84219729; PMID:6328312

A:Reference number: A00641

A:Molecule type: mRNA

A:Residues: 1-1210

A:Cross-references: EMBL:X00588; NID:g31113; PIDN:CAA25240.1; PID:g757924

A:Note: the authors translated the codon AAG for residue 540 as Asn

R:Jishi, S.; Xu, Y.; Stratton, R.H.; Roe, B.A.; Merlino, G.T.; Pastan, I.

Proc. Natl. Acad. Sci. U.S.A. 82, 4920-4924, 1985

A:Title: Characterization and sequence of the promoter region of the human epidermal gtr

A:Accession number: A25772; MUID:85270438; PMID:2991899

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-29 <ISH>

A:Cross-references: GB:M1134; NID:g181981; PIDN:AAA5370.1; PID:g553272

R:Haley, J.; Whittle, N.; Bennett, P.; Kinchington, D.; Ulrich, A.; Waterfield, M.

Oncogene Res. 1, 375-396, 1987

A:Title: The human EGF receptor gene: structure of the 110 kb locus and identification c

A:Accession: A05281

A:Molecule type: protein

A:Residues: 25-30; S', 32-51; 454-467 <WEB>

R:Russo, M.W.; Lukas, T.J.; Cohen, S.; Staros, J.V.

J. Biol. Chem. 260, 5205-5208, 1985

A:Title: Identification of residues in the nucleotide binding site of the epidermal g

A:Reference number: A60143; MUID:85182650; PMID:2985580

A:Accession: A60143

A:Molecule type: protein

A:Residues: 740-744, 'X', 746-747 <RUS>

R:Krocowski, B.; Mosig, G.; Cohen, S.

Nature 309, 270-273, 1984

A:Title: ATP-stimulated interaction between epidermal growth factor receptor and supe

A:Reference number: A38023; MUID:84191554; PMID:6325948

A:Contents: annotation; receptor activity

A:Note: the EGF receptor (and other tyrosine kinases) can nick double-stranded DNA

R:Chen, W.S.; Lazar, C.S.; Lund, K.A.; Welsh, J.B.; Chang, C.P.; Walton, G.M.; Der, C

Cell 59, 33-43, 1989

A:Title: Functional independence of the epidermal growth factor receptor from a domai

A:Reference number: A33331; MUID:90003233; PMID:2790960

A:Contents: annotation; internalization signal

C:Comments: Binding of EGF to the receptor leads to internalization of the EGF-recepto

C:Genetics:

A:Gene: GDB:EGFR

A:Cross-references: GDB:120610; OMIM:131550

A:Map position: 7p12.3-7p12.1

A:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phos

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-120/Product: EGF receptor #status predicted <MAT>

F:25-645/Domain: extracellular #status predicted <EXT>

F:75-300/Domain: EGF receptor extracellular domain repeat <EE1>

F:390-600/Domain: EGF receptor extracellular domain repeat <EE2>

F:646-668/Domain: transmembrane #status predicted <TM>

F:669-1210/Domain: intracellular #status predicted <INT>

F:710-975/Domain: protein kinase homology <KIN>

F:718-726/Region: protein kinase ATP-binding motif

F:999-1046/Region: coated-pit mediated internalization signal

F:1047-1210/Region: inhibitory

F:128,175,352,413,444,528,603/binding site: carbohydrate (Asn) (covalent) #status pre

F:745/Active site: Lys #status experimental

Query Match 13.9%; Score 164; DB 1; Length 1210;

Best Local Similarity 30.5%; Pred. No. 0.0026;

Matches 62; Conservative 24; Mismatches 69; Indels 48; Gaps 11;

Qy 29 SEEPAPNSPLAPSGAGSDVFDGGLGMAKGLDLPDPSPLQRRSEDFVLPSET- 87

Db 1025 SSPSTSRPLSLSLATSN--NSTVACIDRNGLCSCPKEDSFQRRYSSDPTGALTDESI 1082

Qy 88 -DGVVAPLTCSPOPEYVNPQDVRQPPSPRCGLPAPAPGATLEAKTISPGKNGVVKD 146

Db 1083 DDTFL-----PVEEYINQ--SVPRKPGASVONPYHNPOLNP-----APSDPHYQD 1127

Qy 147 VFAFGAGVENEYV--TPQGAAPQHPHPAPFAFDNLYYMDQ-----DP----- 190

Db 1128 --PSTAVAGNEVINTYQ-----PTCVNSFTFSDPAHMAQGSQHSILDNDYQDDE 1176

Qy 191 -PENGAPPSFTKGTPTAENPEYL 212

Db 1177 FPKKAKNGIFKGS-TAENAEYL 1198

RESULT 5

Epidermal growth factor receptor, HER4 - human

C:Species: Homo sapiens (man)

C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999

R:Plowman, G.D.; Culouscou, J.M.; Whittney, G.S.; Green, J.M.; Carlson, G.W.; Foy, L.;

Proc. Natl. Acad. Sci. U.S.A. 90, 1746-1750, 1993

A:Title: Ligand-specific activation of HER4/p180erbB4, a fourth member of the epiderm

A:Reference number: A47253; MUID:93189574; PMID:8383326

A:Accession: A47253
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-1308 <P10>
 A:Cross-references: GB:107868; NID:9337359; PIDN:AAB59446.1; PID:9337360
 A:Note: sequence extracted from NCBI backbone (NCBIP:126842)
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor
 F:716-981/Domain: protein kinase homology <KIN>
 F:724-732/Region: protein kinase ATP-binding motif

Query Match 13.3%; Score 157; DB 2; Length 1308;
 Best Local Similarity 24.2%; Pred. No. 0.0082;
 Matches 62; Conservative 23; Mismatches 75; Indels 96; Gaps 10;

QY 16 RSGGCDLTGLLE-PSEEARSPPLAP-SEGAGSDVEFGDGLGMAKGLQSLPTHPSPQLQ 73
 DB 1067 RDGFAEGAGVSVPRAPRTSTIPKAVAGATATETDSCNCTLRKRVAPHQEDDSSTQ 1126
 QY 74 RYSEDPTVPLPS-----ETDGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPA 126
 DB 1127 RYADPTVFAPEKSPRGELDEEGYMTMRDCKPQOEYLNPE-----ENPFVSR-- 1175
 QY 127 GATLERAKTLSPGKNGVYKVFAGAVENPEYLTPOGGAAPRHPPPA----- 175
 DB 1176 -----KNGDLQ-----ALDNPETHMASNG-----PKADEYVNEPEYL 1209
 QY 176 -----FSPAFDNLVYWDODPPERGA--PSPTEKGTPT----- 205
 DB 1210 NTFANTLGAKEYLKNILMSPEKAKKAFDNPDIWNHSLPRSTLQHDYDQETSTKTFYK 1269
 QY 206 -----AENPEYL 212
 DB 1270 QNGRIRIVANPEYL 1285

RESULT 6

A46302
 P17-associated splicing factor, long form - human
 N:Alternate names: myoblast cell surface antigen 24.1D5; polypyrimidine tract-binding, p17
 C:Species: Homo sapiens (man)
 C:Date: 21-Sep-1993 #sequence revision 18-Nov-1994 #text_change 24-Sep-1999
 C:Accession: A46302; A43557; S29985
 R:Patton, J.G.; Porro, E.B.; Galceran, J.; Tempst, P.; Nadal-Ginard, B.
 Genes Dev. 7, 393-406, 1993
 A:Title: Cloning and characterization of pSF, a novel pre-mRNA splicing factor.
 A:Reference number: A46302; MUID:93194059; PMID:8449401
 A:Accession: A46302
 A:Molecule type: mRNA
 A:Residues: 1-707 <PAT>
 A:Cross-references: EMBL:X70944; NID:938457; PIDN:CAA50283.1; PID:938458
 A:Note: Sequence extracted from NCBI backbone (NCBIP:127206)
 R:Gower, H.J.; Moore, S.E.; Dickson, G.; Elsom, V.L.; Nayak, R.; Walsh, F.S.
 Development 105, 723-731, 1989
 A:Title: Cloning and characterization of a myoblast cell surface antigen defined by 24.1
 A:Reference number: A43557; MUID:90091812; PMID:2480877
 A:Accession: A43557
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 312-707 <GOM>
 A:Cross-references: GB:X16850; NID:923711; PIDN:CAA34747.1; PID:923712
 C:Genetics:
 A:Gene: GDB:SFPO; PSF
 A:Cross-references: GDB:138275
 A:Map position: 4q-4q
 C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleoprotein
 C:Keywords: alternative splicing; pre-mRNA splicing; surface antigen
 F:298-359/Domain: ribonucleoprotein repeat homology <RRM1>
 F:372-438/Domain: ribonucleoprotein repeat homology <RRM2>

Query Match 12.3%; Score 145.5; DB 2; Length 707;
 Best Local Similarity 26.6%; Pred. No. 0.024;
 Matches 58; Conservative 13; Mismatches 82; Indels 65; Gaps 10;

QY 1 GAGGVNHRHRSSSTRSGGDLTLGLEPSEEARSPPLAPSEGAGSDVEFGDGLGMAKGC 60
 DB 11 GGGGGGHR-----GGGGGCGLHDFR-----SPPMGLNONGNPGPGGGS 54
 QY 61 LQSLPHDPSPLQRYSE-DPTVPLPSTGCVAPLTCSPQPEYVNOQDVRPQPPSPRE-- 117
 DB 55 GPKRPITPPPPHQQOQPPQOQPPPPHQP--PHQPHQQOQP--PPPPDSS 107
 QY 118 -----GPLPARPAGATLERAKTLSPGKNGVYKVFAGAVENPEYLTPOGGAAPQ 169
 DB 108 KPVVAGCGPAPGVGAPRPAASSAPRATPTPTSA-----PPGSGRPT 150
 QY 170 PHPPAFSPAEDNLVYWDODPPERGAPST--FKGTPT 205
 DB 151 PRPPPAVTS-----PP--GAPPPPPSSGVP 176

RESULT 7

TVEVLV
 Protein-tyrosine kinase (EC 2.7.1.112) erbB - avian leukosis virus
 N:Contains: amino end of gag protein; env protein fragment; protein-tyrosine kinase
 C:Species: avian leukosis virus, ALV
 C:Date: 31-Dec-1991 #sequence revision 31-Dec-1991 #text_change 11-Jun-1999
 C:Accession: B00643; A00643
 R:Nilsson, T.W.; Maroney, P.A.; Goodwin, R.G.; Rottman, F.M.; Crittenden, L.B.; Ralnes
 Cell 41, 719-726, 1985
 A:Title: c-erbB activation in ALV-induced erythroblastosis: novel RNA processing and
 A:Reference number: A00643; MUID:85228222; PMID:2988784
 A:Accession: B00643
 A:Molecule type: mRNA
 A:Residues: 1-698 <N1L>
 A:Cross-references: GB:M10066; GB:M13881; NID:9211749; PIDN:AAA48763.1; PID:9211750
 A:Note: In Genbank entry CHKERBBF, release 109.0, the source is designated as Gallus
 C:Comment: This protein is synthesized as a gag-env-erbB protein.
 C:Genetics:
 A:Gene: gag-env-erbB
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; oncogene; phosphotransferase; transforming protein; tyrosine-specific
 F:1-6/Product: gag protein (fragment) #status predicted <GAG>
 F:7-59/Product: env protein (fragment) #status predicted <ENV>
 F:60-698/Product: protein-tyrosine kinase erbB #status predicted <ERB>
 F:194-459/Domain: protein kinase homology <KIN>
 F:202-210/Region: protein kinase ATP-binding motif
 F:229/Active site: Lys #status predicted

Query Match 11.7%; Score 138; DB 1; Length 698;
 Best Local Similarity 24.5%; Pred. No. 0.073;
 Matches 54; Conservative 25; Mismatches 59; Indels 62; Gaps 10;

QY 35 RSPV-----APSEGAGSDVEFGDGLGMAKGLQSLPTHPSPQLQRYSEDPVPLPSET-- 87
 DB 515 RTFLLSLSLSTSNNSATNCID-----RNGCGHPRFEDSFVQRSSDFTGNFLESID 566
 QY 88 DGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPAGATLERAKTLSPGKNGVYKQV 147
 DB 567 DGFL-----PAPEYVNO--LMPKPS-----TAAVNOQYNNI 597
 QY 148 F-----AFGAVENPEYLTPOGGAAPRHPPAFSPAEDNLVYWDQ----- 188
 DB 598 SLTAISKLPMDSKYQNSHSTAVIDNPEYL-----NTNQSPLAKYFESSPIYIOSGNH 649
 QY 189 ----DPPE-----RGAPPSTFKGTPTAENPEYLGADV 217
 DB 650 QINLDNDPDYQDPLPNETKPNGLKTVPAENPEYLRVAAP 669

RESULT 8

TYCVHLV
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) erbB
 C:Species: Gallus gallus (chicken)
 C:Date: 28-Feb-1986 #sequence revision 05-May-1995 #text_change 04-Feb-2000

A:Experimental source: subject C.J. (large allele)
A:Accession: C40750
A:Molecule type: DNA
A:Residues: 35-127, 'R', 129-148, 'R', 150-151, 153-187, 'K', 189-272, 'S', 274-336, 'S', 338-339
A:Cross-references: GB:562929
A:Experimental source: subject M.V.O. (large allele)
A:Accession: A40750
A:Molecule type: DNA
A:Residues: 35-183, 245-270, 'Q', 272-392 <AZ>
A:Cross-references: GB:562928
A:Experimental source: subject C.J. (medium allele)
A:Note: authors translated the codon CAA for residue 272 as Arg
R:Meda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985
A:Title: Differential RNA splicing and post-translational cleavages in the human salivary
A:Reference number: A92492; MUID:85289325; PMID:2993301
A:Accession: C25372
A:Molecule type: mRNA
A:Residues: 1-183, 245-392 <MAE>
A:Cross-references: GB:603204; NID:g190485; PIND:AAA60185.1; PID:g190486
A:Note: alternatively splice forms lacking portions of the repeat region were also for
R:Lyons, K.M.; Stein, J.H.; Smithies, O.
Genetics 120, 267-278, 1988
A:Title: Length polymorphisms in human prolactin-rich protein genes generated by Interg
A:Reference number: 502127; MUID:89121440; PMID:2851479
A:Accession: 502128
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 35-127, 250-273, 'R', 275-277, 'R', 279-336, 'S', 338-392 <LYO>
A:Cross-references: EMBL:X07517
A:Accession: 502127
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 35-183, 245-392 <LY2>
A:Cross-references: EMBL:X07516
R:Kauffman, D.; Hofmann, T.; Bennick, A.; Keller, P.
Biochemistry 25, 2387-2392, 1986
A:Title: Basic prolactin-rich proteins from human parotid saliva: complete covalent str
A:Reference number: A90502; MUID:86243355; PMID:3521730
A:Accession: A03293
A:Molecule type: protein
A:Residues: 17-38, 'AP', 41-51, 92-148, 'R', 150-152 <KA2>
A:Note: among nine basic prolactin-rich peptides isolated from the saliva, this peptide
A:Accession: A90502
A:Molecule type: protein
A:Residues: 275-336, 'S', 338-392 <KAU>
R:Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 94, 1991-1999, 1983
A:Title: Further fractionation of basic prolactin-rich peptides from human parotid saliv
A:Reference number: A91974; MUID:84161824; PMID:6671974
A:Contents: P-H
A:Accession: A91974
A:Molecule type: protein
A:Residues: 'S', 338-392 <SAI>
A:Azen, E.; Lyons, R.; McConigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Van
Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984
A:Reference number: A94003; MUID:84298176; PMID:6089212
A:Accession: A05261
A:Molecule type: DNA
A:Residues: 35-39, 'P', 41-84, 'G', 86, 'R', 87-154, 'R', 218-246; 300-306, 'T', 308-329, 'C', 331D
A:Accession: A05262
A:Molecule type: DNA
A:Residues: 'N', 57-59, 'A', 61-69; 334-336, 'S', 338-339, 'R', 341-392 <AZ5>
R:Kauffman, D.; Wong, R.; Bennick, A.; Keller, P.
Biochemistry 21, 6558-6562, 1982
A:Title: Basic prolactin-rich proteins from human parotid saliva: complete covalent str
A:Reference number: A90464; MUID:83101329; PMID:6924859
A:Contents: IB-9
A:Accession: A90464
A:Molecule type: protein
A:Residues: 92-127, 'R', 129-148, 'R', 150-152 <KA3>
R:Isemura, S.; Saitoh, E.; Sanada, K.
J. Biochem. 91, 2067-2075, 1982

A:Cross-references: EMBL:X68249; NID:g64955; PID:g64956
C:Superfamily: proline-rich protein

Query Match 11.4%; Score 134.5; DB 2; Length 379;
Best Local Similarity 28.0%; Pred. No. 0.064;
Matches 59; Conservative 16; Mismatches 97; Indels 39; Gaps 9;

QY 2 AGCVHHRHSSSTRSGGDLTLGLEPSEEARSPPLAPSEG---AGSDVFDGLMGAA 58
DB 134 SGGPRHDDSDSGNRSSPREVGRHRSPLDLSPPSTSGMKHSSSAPPSPGRQA 193
QY 59 KGLQSLPTHPSPLOQRYSEDPVPLPSETDGYAPLTCSPQPEYVNDVAPRPPSP--- 115
DB 194 -GAPPAASQNAKKRYNR--EKPLPPTPGHRAAPVAPVPPSPINSRSPSAHSQPPPPPY 250
QY 116 -----REGPLPAARAGATLERAKTLSPGKNGVAKDVFAGAVENPE--YITPGGA 166
DB 251 RQPTSLNGPPSPINPEAPDELPOHNSLHKRTAGPVR---GLAPPQSVHLSP--GGN 304
QY 167 APQHPAPAFSPADNLYWDDPPERGAPP 197
DB 305 RP---PPPA-----KQPPGKGAAP 320

RESULT 13
T29293
hypothetical protein C50F7.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T29293
R:Johnson, D.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C50F7.
A:Reference number: Z20601

A:Accession: T29293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-309 <JOH>
A:Cross-references: EMBL:U41557; PID:AAA83301.1; CESP:C50F7.5
C:Genetics:
A:Gene: CESP:C50F7.5
C:Superfamily: unassigned collagens

Query Match 11.3%; Score 134; DB 2; Length 309;
Best Local Similarity 24.2%; Pred. No. 0.055;
Matches 54; Conservative 25; Mismatches 74; Indels 70; Gaps 10;

QY 7 HHRHSSSTRSGGDLTLGLEPSEEARSPPLAPSEGSDVFDGLMGAAKGLQSLPT 66
DB 31 HHHHKTAPRTSRGIATTTTFAPTSSDL---PIAGSSSA----- 65
QY 67 HDPSPLDRYSPD---TYPLPSE--TDGVVAALTCSPQPEYVNDVAPRPPSPREGP 119
DB 66 ----PVLAASADPLPTSVVPQPSNEPSPGVAP---SDSPSPGPPS--PGVNPSEDP 116
QY 120 LPAARAGATLERAKTLSPGKNGVAKDVFAGAVENPEYITPGGAAPQ---HPP 174
DB 117 QPSPGPSPGPVDPSEDPQ-----SVPSSEDPQSPGPPSPGVDSEDDPQ 162
QY 175 AFSAPFNLYWDDPPERGAPSTFGTPPAENPEYGLDVP 217
DB 163 SVPSSED-----PQSPGPPSPGPVDP--SEDPQPSGSSSP 195

RESULT 14
T05264
probable serine/threonine-specific protein kinase (EC 2.7.1.-) T4L20.20 - Arabidopsis th
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000
C:Accession: T05264
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysnaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, September 1998

A:Reference number: Z15406
A:Accession: T05264
A:Molecule type: DNA
A:Residues: 1-674 <BEV>
A:Cross-references: EMBL:AL023094
A:Experimental source: cultivar Columbia; BAC clone T4L20
C:Genetics:
A:Map position: 4
A:Introns: 180/1; 221/1; 381/1; 421/1; 444/3; 470/2; 518/3; 583/3
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: phosphotransferase; protein kinase

Query Match 11.3%; Score 134; DB 2; Length 674;
Best Local Similarity 23.9%; Pred. No. 0.13;
Matches 59; Conservative 22; Mismatches 76; Indels 90; Gaps 12;

QY 17 SGGDLTLGLEPSEEARSPPLAPSEGSDVFDGLMGAAKGLQSLPTHD----- 68
DB 15 SNGTPSNGTSPSNSSPPRP--PSSPPSSI-----SAPPDIGASFSPP 58
QY 69 PSLQRYSEDPVPLPSETDGYAPLTCSPQPEYVNDVAPRPPSPREG-----PLPA 122
DB 59 PAPPTQETSPSTSPSSS-----PPVYANSP---QTPENPSPAP--EGSTPYTPPAP 107
QY 123 ARPAGATLERAKTLSPGKNGVAKDVFAGAVENPEYITPO-----GGAPOHPHP 174
DB 108 QTPSNQSPERTTPSPGAN---DDRNTNGNNRNDSTSPSSGRTSDGSPSPPR 164
QY 175 AF-----SPAFNLYWDDPPERGAPST- 199
DB 165 SISPPQNSGDSSSGLLLLAVCICICNKKKKKKQVHHMYNNN--PYGAPSGNG 223
QY 200 --FKGTP 204
DB 224 GYKGTGTP 230

RESULT 15
A53183
epidermal growth factor receptor precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: A53183; A43818; S24942; A28941; S45325; I49643
R:Luetteke, N.C.; Phillips, H.K.; Qiu, T.H.; Copeland, N.G.; Earp, H.S.; Jenkins, N.A
Genes Dev. 8, 399-413, 1994
A:Title: The mouse waved-2 phenotype results from a point mutation in the EGF recepto
A:Reference number: A53183; MUID:94170986; PMID:8125255
A:Accession: A53183
A:Molecule type: mRNA
A:Residues: 1-1210 <LUET>
A:Cross-references: GB:U03425
R:Aviv, A.; Lax, I.; Ullrich, A.; Schlessinger, J.; Givol, D.; Morse, B.
Oncogene 6, 673-676, 1991
A:Title: Comparison of EGF receptor sequences as a guide to study the ligand binding
A:Reference number: A43818; MUID:91232866; PMID:2030916
A:Accession: A43818
A:Molecule type: mRNA
A:Residues: 1-714 <AVI>
R:Eisinger, D.P.; Serrero, G.
submitted to the EMBL Data Library, June 1992
A:Reference number: S24942
A:Accession: S24942
A:Molecule type: mRNA
A:Residues: 969-971, 'K', 973-1115, 'D' <EIS>
A:Cross-references: EMBL:Z12608
R:Heisermann, G.J.; Gili, G.N.
J. Biol. Chem. 263, 13152-13159, 1988
A:Title: Epidermal growth factor receptor threonine and serine residues phosphorylate
A:Reference number: A28941; MUID:88350814; PMID:3138223
A:Accession: A28941
A:Molecule type: protein

A:Residues: 689-694,'X',696-704,'L',706-707;989-992,'XX',995-996,'X',998-1000;1002-1009,
 R:Hiibs, M.L.; Dunn, A.R.; Alexander, W.S.
 submitted to the EMBL Data Library, April 1994
 A:Description: The complete cDNA sequence of the Mouse Epidermal Growth Factor Receptor
 A:Reference number: S45325
 A:Accession: S45325
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-971,'K',973-1210 <VER>
 A:Cross-references: EMBL:X78987; NID:q488830; PIDN:CAA55587.1; PID:q488831
 R:Parls, B.C.; Das, S.K.; Andrews, G.K.; Day, S.K.
 Proc. Natl. Acad. Sci. U.S.A. 90, 55-59, 1993
 A>Title: Expression of the epidermal growth factor receptor gene is regulated in mouse B
 A:Reference number: 149643; MUID:93126380; PMID:7678348
 A:Accession: 149643
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 12-20,22-132 <RES>
 A:Cross-references: GB:L06864; NID:q193001; PIDN:AAA53029.1; PID:q567201
 C:Genetics:
 A:Gene: EGFR
 C:Superfamily: epidermal growth factor receptor; protein kinase homology
 C:Keywords: ATP; growth factor receptor; kinase-related transforming protein; phosphop
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:648-670/Domain: transmembrane #status predicted <TM>
 F:712-977/Domain: protein kinase homology <KIN>
 F:720-728/Region: protein kinase ATP-binding motif
 F:680-695/Binding site: phosphate (Thr) (covalent) #status experimental
 F:697-1070/Binding site: phosphate (Ser) (covalent) #status experimental
 F:993/Binding site: (or 997) phosphate (Ser) (covalent) #status experimental
 F:1028/Binding site: (or 1030 or 1032) phosphate (Ser) (covalent) #status experimental
 F:1197/Binding site: phosphate (Tyr) (covalent) #status experimental

Query Match 11.3%; Score 134; DB 2; Length 1210;
 Best local Similarity 26.4%; Pred. No. 0.24;
 Matches 58; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

QY 47 DVFDGLGMAKGLQSLPTHDPSP-----LQRY 76
 ||| : : : : : ||||
 DB 1011 DVYDADEYLIPQGFNFSSTSRTPLLSLATSNSTVACINRNGSCRYKEDAFLOQYS 1070
 QY 77 EDPTVPLPSET--DGYVAPLTCSPQPEYVQNPVRRQPPSPREGPLPAARRAGATLERAK 134
 ||| : : : : : ||||
 DB 1071 SDPTGAVTEENIDDAFL-----PVPEYVQ--SVPRKPGSVQNPVYHNOPLHP----- 1117
 QY 135 TLPGRKNGVVKDYAFAGAVENPEYL--TPQGAAPQPHPPAFSPAFDNLXYWDQ----- 188
 :||: : : ||||| : : : : :
 DB 1118 --APGRDLHQN--PHSNVGNPEYLNTAQ-----PTCLSSGFNSPALMIQKSHQ 1164
 QY 189 ----DP-----PERGAPPSTFKGPTAENPEYLGIDVP 217
 : : : : : ||||| : : : : :
 DB 1165 MSIDNPDYOODFFPKETKPKNGIFKG--PTAENAEYLVAPP 1203

Search completed: April 28, 2003, 13:42:33
 Job time : 11.9833 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:08 ; Search time 17.4181 Seconds
(without alignments)
1660.081 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182
Sequence: 1 GACGMVHRRSRSSTRSGC.....STFKGPTAEPEYLGDVP 217

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182	100.0	217	22	AAE13121
2	1182	100.0	397	22	AAE13122
3	1182	100.0	555	22	AAE13108
4	1182	100.0	564	22	AAE13110
5	1182	100.0	690	22	AAE13109
6	1182	100.0	697	22	AAE13111
7	1182	100.0	1179	22	AAE13123
8	1182	100.0	1223	23	AAU98923
9	1182	100.0	1255	21	AAV92620
10	1182	100.0	1255	22	AAE12130

11	1182	100.0	1255	22	AAE60167
12	1182	100.0	1255	23	AAU74545
13	1177	99.6	266	21	AAE21201
14	1177	99.6	266	23	AAE51146
15	1177	99.6	583	23	AAE20483
16	1177	99.6	587	23	AAE20481
17	1177	99.6	589	23	AAE20484
18	1177	99.6	600	23	AAE20482
19	1177	99.6	919	21	AAE21203
20	1177	99.6	919	23	AAE51148
21	1177	99.6	1255	17	AAW01111
22	1177	99.6	1255	20	AAW92406
23	1177	99.6	1255	21	AAE21158
24	1177	99.6	1255	21	AAV84780
25	1177	99.6	1255	22	AAE85458
26	1177	99.6	1255	22	AAE88267
27	1177	99.6	1255	23	AAE20479
28	1177	99.6	1255	23	AAE20479
29	1177	99.6	1255	23	AAE51143
30	1177	99.6	1255	23	AAU77114
31	1177	99.6	1433	19	AAE39568
32	1086	91.9	287	14	AAE53133
33	961	81.3	920	23	AAE51152
34	961	81.3	926	23	AAE51153
35	961	81.3	1256	21	AAE21206
36	961	81.3	1256	22	AAE62860
37	961	81.3	1256	23	AAE51151
38	954	80.7	1256	21	AAE21159
39	954	80.7	1256	23	AAE51144
40	864	73.1	1200	21	AAE21208
41	711	60.2	293	21	AAE58875
42	648.5	54.9	135	21	AAE43787
43	164	13.9	1210	21	AAE19259
44	164	13.9	1210	21	AAE50616
45	164	13.9	1210	22	AAE68420

ALIGNMENTS

```
RESULT 1
AAE13121
ID AAE13121 standard; Protein; 217 AA.
AC AAE13121;
XX 28-JAN-2002 (first entry)
DE Mature human HER-2 membrane distal intracellular domain.
EE
DE Immunostimulatory fusion protein; IFP; antigen component; therapy.
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW human; HER-2 membrane distal intracellular domain.
XX Homo sapiens.
XX WO200174855-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WC-US10515.
XX 30-MAR-2000; 2000US-193504P.
XX (DEND-) DENDREON CORP.
XX Laus R, Vidovic D, Graddis T;
XX WPT: 2001-662965/76.
XX N-PSDB: AAD21572.
XX An immunostimulatory fusion protein comprising the intracellular domain
```

HER2 transgene pla
Human HER2 (ErbB2)
Human HER-2/neu pr
Human HER-2/neu on
Human HER-2/neu on
Human protein for
Human protein for
Human protein for
Human protein for
Human protein for
HER-2/neu extracel
HER-2/neu protein.
Human HER-2/neu on
Human HER-2/neu pr
Amino acid sequenc
Human HER-2/neu pr
HER2/neu amino aci
Human Her-2 protei
Human Her-2/neu pr
Human Her-2/neu on
Human Her-2/neu po
Sequence of c-erbB
HER-2/neu polypept
Human Her-2/neu ex
Mouse Her-2/neu ex
Mouse Her-2/neu pr
Amino acid sequenc
Mouse Her-2/neu on
Rat Her-2/neu prot
Rat Her-2/neu onco
Human HER-2/neu pr
Breast and ovarian
Human cancer assoc
Amino acid sequenc
Human BCF receptor
Amino acid sequenc

PT of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 XX
 PS Claim 2; Page 31; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is
 CC mature human HER-2 membrane distal intracellular domain.
 CC This sequence is used in the HER500 and HER300 GM-CSF fusion constructs
 CC of the invention.

SQ Sequence 217 AA;

Query Match 100.0%; Score 1182; DB 22; Length 217;

Best Local Similarity 100.0%; Pred. No. 1.8e-91;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGCMVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVFDGDLGMGAAG 60
 DB 1 GAGCMVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVFDGDLGMGAAG 60
 OY 61 LQSLPTHPDPSLDRYSDDPVPPLPSETDGVVAPLTCSPOEYVNOPVRRQPPSPREGPL 120
 DB 61 LQSLPTHPDPSLDRYSDDPVPPLPSETDGVVAPLTCSPOEYVNOPVRRQPPSPREGPL 120
 OY 121 PAARPGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPPPAPSPAF 180
 DB 121 PAARPGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPPPAPSPAF 180
 OY 181 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217
 DB 181 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217

RESULT 2

AAE13122 standard; Protein; 397 AA.

XX AAE13122;

DT 28-JAN-2002 (first entry)

DE NY-ESO-1C-HER-2 membrane distal intracellular domain fusion protein.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW human; HER-2 membrane distal intracellular domain; NY-ESO-1C;
 KW autologous cancer/testis antigen.

OS Homo sapiens.

PN WO200174855-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US10515.

PR 30-MAR-2000; 2000US-193504P.

PA (DEND-) DENDREON CORP.

PI Laus R, Vidovic D, Graddis T;

DR WPI: 2001-662965/76.

XX N-PSDB; AAD21573.

PT An immunostimulatory fusion protein comprising the intracellular domain
 PT of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 XX
 PS Disclosure; Page 54-55; 59pp; English.

CC The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is a
 CC fusion protein which comprises human autologous
 CC cancer/testis antigen, NY-ESO-1C and mature human HER-2 membrane distal
 CC intracellular domain.

SQ Sequence 397 AA;

Query Match 100.0%; Score 1182; DB 22; Length 397;

Best Local Similarity 100.0%; Pred. No. 3.6e-91;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGCMVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVFDGDLGMGAAG 60
 DB 181 GAGCMVHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVFDGDLGMGAAG 240
 OY 61 LQSLPTHPDPSLDRYSDDPVPPLPSETDGVVAPLTCSPOEYVNOPVRRQPPSPREGPL 120
 DB 241 LQSLPTHPDPSLDRYSDDPVPPLPSETDGVVAPLTCSPOEYVNOPVRRQPPSPREGPL 300
 OY 121 PAARPGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPPPAPSPAF 180
 DB 301 PAARPGATLERAKTLSPKNGVYKDVFAFGAVENPEYLTPOGAAPQHPPPAPSPAF 360
 OY 181 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217
 DB 361 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 397

RESULT 3

AAE13108 standard; Protein; 555 AA.

XX AAE13108;

DT 28-JAN-2002 (first entry)

DE Human HER500 fusion protein construct.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal intracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human;
 KW HER-2 protein; HER500 fusion protein.

OS ChimERIC - Homo sapiens.

PN WO200174855-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US10515.

PR 30-MAR-2000; 2000US-193504P.

PA (DEND-) DENDREON CORP.

PI Laus R, Vidovic D, Graddis T;

XX

DR WPI: 2001-662965/76.
DR N-PSDB: AAD21564.
PT An immunostimulatory fusion protein comprising the intracellular domain
of HER-2 and an antigen elicits an immune response to the antigen and
is useful for the treatment of associated cancer associated -
XX
XX
PS Claim 7; Page 26; 59pp: English.
XX
CC The invention relates to immunostimulatory fusion proteins (IFP) and
nucleic acid molecules encoding such proteins. The IFPs comprise a
polypeptide antigen component and an immunostimulatory component derived
from the intracellular domain of HER-2 protein which is effective to
elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
immune response to the antigen. IFP or superactivated dendritic cells
are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
associated with a particularly antigen. The present sequence is HER500
fusion protein construct which comprises human PAP signal
sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
sequence, mature HER-2 membrane distal extracellular and intracellular
domains and a C-terminal tag.
XX
SQ Sequence 555 AA;
Query Match 100.0%; Score 1182; DB 22; Length 555;
Best Local Similarity 100.0%; Pred. No. 5,3e-91;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGMVHHRHSSSTRSGGDLTGLPESEEARSPPLASBEGAGSDVFDGLMGAAKG 60
DB 330 GAGGMVHHRHSSSTRSGGDLTGLPESEEARSPPLASBEGAGSDVFDGLMGAAKG 389
QY 61 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRGSL 120
DB 390 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRGSL 449
QY 121 PAARPAGATLERAKTSLSPKNGVYKDYAFAGAVENPEYLPQGAAPQHPPPAFSPAF 180
DB 450 PAARPAGATLERAKTSLSPKNGVYKDYAFAGAVENPEYLPQGAAPQHPPPAFSPAF 509
QY 181 DNLTYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217
DB 510 DNLTYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 546
RESULT 4
ID AAE13110 standard; Protein: 564 AA.
XX
XX AAE13110;
AC
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.
XX
XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
immunostimulatory component; T-cell mediated immune response; DC;
dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
PAP protein; Ala Arg linker; membrane distal extracellular domain;
membrane distal intracellular domain; C-terminal tag; human; OVA;
HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Unidentified.
XX
XX
PN WO200174855-A2.
PD 11-OCT-2001.
XX
XX
PF 30-MAR-2001; 2001MO-US10515.
XX
XX 30-MAR-2000; 2000US-193504P.
XX

PA (DEND-) DENDREON CORP.
XX
PI Laus R, Vidovic D, Graddis T;
XX
XX WPI: 2001-662965/76.
DR N-PSDB: AAD21566.
XX
XX
PT An immunostimulatory fusion protein comprising the intracellular domain
of HER-2 and an antigen elicits an immune response to the antigen and
is useful for the treatment of associated cancer associated -
XX
XX
PS Claim 7; Page 26; 59pp: English.
XX
CC The invention relates to immunostimulatory fusion proteins (IFP) and
nucleic acid molecules encoding such proteins. The IFPs comprise a
polypeptide antigen component and an immunostimulatory component derived
from the intracellular domain of HER-2 protein which is effective to
elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
immune response to the antigen. IFP or superactivated dendritic cells
are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
associated with a particularly antigen. The present sequence is HER500
fusion protein construct which comprises human PAP signal
sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
sequence, mature HER-2 membrane distal extracellular domain, an
Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
HER-2 membrane distal intracellular domain and a C-terminal tag.
XX
SQ Sequence 564 AA;
Query Match 100.0%; Score 1182; DB 22; Length 564;
Best Local Similarity 100.0%; Pred. No. 5,4e-91;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGMVHHRHSSSTRSGGDLTGLPESEEARSPPLASBEGAGSDVFDGLMGAAKG 60.
DB 339 GAGGMVHHRHSSSTRSGGDLTGLPESEEARSPPLASBEGAGSDVFDGLMGAAKG 398
QY 61 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRGSL 120
DB 399 LQSLPTDPSFLQRYSEDPYPLPSETDGYAPLTCSPOPEYVQPDVRRPPSPRGSL 458
QY 121 PAARPAGATLERAKTSLSPKNGVYKDYAFAGAVENPEYLPQGAAPQHPPPAFSPAF 180
DB 459 PAARPAGATLERAKTSLSPKNGVYKDYAFAGAVENPEYLPQGAAPQHPPPAFSPAF 518
QY 181 DNLTYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217
DB 519 DNLTYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 555
RESULT 5
ID AAE13109 standard; Protein: 690 AA.
XX
XX AAE13109;
AC
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500-hGM-CSF fusion protein construct.
XX
XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
immunostimulatory component; T-cell mediated immune response; DC;
dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
PAP protein; Ala Arg linker; membrane distal extracellular domain;
membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
HER-2 protein; granulocyte-macrophage colony stimulating factor;
HER500-hGM-CSF fusion protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Synthetic.
XX
XX
PN WO200174855-A2.
XX

PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US10515.
XX
XX 30-MAR-2000; 2000US-193504P.
XX
XX (DEND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddis T;
XX
XX WPI: 2001-662965/76.
DR N-PSDB; AAD21565.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and
PT is useful for the treatment of associated cancer associated -
XX
XX
PS Claim 7; Page 26; 59pp: English.
XX
XX The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFP or superactivated dendritic cells
CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC hgm-CSF fusion protein construct which comprises human PAP
CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
CC signal sequence, mature HER-2 membrane distal extracellular and
CC intracellular domains, an Ala Ala linker, a mature human granulocyte-
CC macrophage colony stimulating factor (GM-CSF) sequence and a
CC C-terminal tag.
XX
XX
SQ Sequence 690 AA:
Query Match 100.0%; Score 1182; DB 22; Length 690;
Best Local Similarity 100.0%; Pred. No. 6.8e-91;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFPGDLGMAAKG 60
DB 330 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFPGDLGMAAKG 389
OY 61 LQSLPTHDPSPLQRYSDPVPPLPSETDGYVAALTCSPQPEYVNOQPVRRQPPSPREGPL 120
DB 390 LQSLPTHDPSPLQRYSDPVPPLPSETDGYVAALTCSPQPEYVNOQPVRRQPPSPREGPL 449
OY 121 PAARPAGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180
DB 450 PAARPAGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 509
OY 181 DNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217
DB 510 DNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 546
RESULT 6
AAE13111 standard; Protein: 697 AA.
XX
XX AAE13111:
XX
XX 28-JAN-2002 (first entry)
XX
XX Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.
XX
XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;

KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.
XX
XX Chimeric - Homo sapiens.
OS Chimeric - Rattus norvegicus.
OS Chimeric - Unidentified.
XX
XX WO200174855-A2.
XX
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001MO-US10515.
XX
XX 30-MAR-2000; 2000US-193504P.
XX
XX (DEND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddis T;
XX
XX WPI: 2001-662965/76.
DR N-PSDB; AAD21567.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and
PT is useful for the treatment of associated cancer associated -
XX
XX
PS Claim 7; Page 27; 59pp: English.
XX
XX The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFP or superactivated dendritic cells
CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer.
CC associated with a particularly antigen. The present sequence is HER500
CC rGM-CSF fusion protein construct which comprises human PAP
CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
CC signal sequence, mature HER-2 membrane distal extracellular domain,
CC an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide,
CC HER-2 membrane distal intracellular domain, an Ala Ala linker, a mature
CC rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence
CC and a C-terminal tag.
XX
XX
SQ Sequence 697 AA:
Query Match 100.0%; Score 1182; DB 22; Length 697;
Best Local Similarity 100.0%; Pred. No. 6.9e-91;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFPGDLGMAAKG 60
DB 339 GAGGMVHHRSSSTRSGGDLTLGLEPSEEAARSPRLASEGAGSDVFPGDLGMAAKG 398
OY 61 LQSLPTHDPSPLQRYSDPVPPLPSETDGYVAALTCSPQPEYVNOQPVRRQPPSPREGPL 120
DB 399 LQSLPTHDPSPLQRYSDPVPPLPSETDGYVAALTCSPQPEYVNOQPVRRQPPSPREGPL 448
OY 121 PAARPAGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180
DB 459 PAARPAGATLERAKTILSPKNGVYKDFAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 518
OY 181 DNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 217
DB 519 DNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 555
RESULT 7
AAE13123
ID AAE13123 standard; Protein: 1179 AA.
XX
XX AAE13123:
XX
XX 28-JAN-2002 (first entry)

XX DE Human SART3-IC-HER-2 membrane distal intracellular domain fusion protein.
 XX XX Immunostimulatory fusion protein: IFP; antigen component; therapy;
 KW immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW human; HER-2 membrane distal intracellular domain; SART3-IC;
 KW squamous cell carcinoma antigen; fusion protein.
 XX OS Homo sapiens.
 XX OS WO200174855-A2.
 XX PN 11-OCT-2001.
 XX PD 30-MAR-2001; 2001WO-US10515.
 XX PE 30-MAR-2000; 2000US-193504P.
 XX PR 30-MAR-2000; 2000US-193504P.
 XX PA (DEND-) DENDREON CORP.
 XX PI Laus R, Vidovic D, Graddis T;
 XX DR WPI: 2001-662965/76.
 XX DR N-PSDB: AAD21574.
 XX XX An immunostimulatory fusion protein comprising the intracellular domain
 PT of HER-2 and an antigen elicits an immune response to the antigen and
 PT is useful for the treatment of associated cancer associated -
 XX XX
 XX XX Disclosure: Page 55-58; 59pp; English.
 XX PS The invention relates to immunostimulatory fusion proteins (IFP) and
 CC nucleic acid molecules encoding such proteins. The IFPs comprise a
 CC polypeptide antigen component and an immunostimulatory component derived
 CC from the intracellular domain of HER-2 protein which is effective to
 CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 CC immune response to the antigen. IFP or superactivated dendritic cells
 CC are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is a
 CC fusion protein which comprises human squamous cell
 CC carcinoma antigen, SART3-IC and mature human HER-2 membrane distal
 CC intracellular domain.
 CC XX
 SO Sequence 1179 AA;
 Query Match 100.0%; Score 1182; DB 22; Length 1179;
 Best Local Similarity 100.0%; Pred. No. 1.2e-90;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GAGGWHHRHSSSTRSGGDLTLGLEPSEEARPSPLAPSBGASDVFDGDLGGAAG 60
 Db 963 GAGGWHHRHSSSTRSGGDLTLGLEPSEEARPSPLAPSBGASDVFDGDLGGAAG 1022
 Oy 61 LOSLPTHPSPLORYSEDPYPLPSETDGYVAPLTCSPOPEYVNPDPVPPSPREGPL 120
 Db 1023 LOSLPTHPSPLORYSEDPYPLPSETDGYVAPLTCSPOPEYVNPDPVPPSPREGPL 1082
 Oy 121 PAARAGATLERAKTILSPGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180
 Db 1083 PAARAGATLERAKTILSPGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 1142
 Oy 181 DNLTYWDDPPERGAPSTFKGTPTAENPEYLTGLDVP 217
 Db 1143 DNLTYWDDPPERGAPSTFKGTPTAENPEYLTGLDVP 1179

RESULT 8
 AAU98923
 ID AAU98923 standard; Protein: 1223 AA.
 XX AC AAU98923;
 XX XX

DT 24-SEP-2002 (first entry)
 XX DE Human breast cancer antigen, Her2 variant.
 XX XX Human; Her2; cyrostatic; antiviral; immunostimulant;
 KW cell-mediated immune response; tumour; breast cancer;
 KW virus infection; prostate cancer; colorectal cancer; pancreatic cancer;
 KW lymphoma; leukaemia; hepatitisvirus; lentivirus; herpesvirus;
 KW human immunodeficiency virus; HIV; flavivirus; pestivirus.
 XX OS Homo sapiens.
 XX OS WO200240059-A2.
 XX PN 23-MAY-2002.
 XX PD 01-NOV-2001; 2001WO-US45626.
 XX PE 01-NOV-2000; 2000US-0704232.
 XX PR 01-NOV-2000; 2000US-0704232.
 XX PA (AMBI-) AMERICAN FOUND BIOLOGICAL RES INC.
 XX PA (MINC/) MINCHEFF M S.
 XX PA (LOUK/) LOUKINOV D I.
 XX PA (ZOUB/) ZOUBAK S.
 XX PI Mincheff MS, Loukinov DI, Zoubak S;
 XX DR WPI: 2002-527524/56.
 XX DR N-PSDB: ABK86207.
 XX XX
 PT Inducing a cell-mediated immune response against a target antigen, by
 PT reducing undesired cells and stimulating presentation of an antigen by
 PT a cell, comprises administering a polynucleotide encoding a variant of
 PT an antigen -
 XX XX
 XX PS Claim 21; Page 134-138; 146pp; English.
 CC The invention relates to a method of inducing a cell-mediated immune
 CC response against a cell comprising a target antigen (I) in a subject,
 CC treating a subject having undesired cells, for example tumour cells
 CC or virally infected cells (C), reducing the number of (C) in a subject,
 CC and stimulating presentation of (I) by a cell. This is done by
 CC administering a polynucleotide (II) encoding a variant of (I), so that
 CC (II) expressed in a cell and cell-mediated immune response is induced.
 CC The method can be used to treat prostate cancer, breast cancer,
 CC colorectal cancer and pancreatic cancer, as well as lymphomas and
 CC leukaemias. The method is also useful in treating chronic viral
 CC infections such as those caused by hepatitisviruses, lentiviruses
 CC (including human immunodeficiency virus (HIV)), herpesviruses and the
 CC human breast cancer antigen, Her2 variant, used as a target
 CC antigen in the method of the invention.
 CC XX
 SO Sequence 1223 AA;
 Query Match 100.0%; Score 1182; DB 23; Length 1223;
 Best Local Similarity 100.0%; Pred. No. 1.3e-90;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GAGGWHHRHSSSTRSGGDLTLGLEPSEEARPSPLAPSBGASDVFDGDLGGAAG 60
 Db 1006 GAGGWHHRHSSSTRSGGDLTLGLEPSEEARPSPLAPSBGASDVFDGDLGGAAG 1065
 Oy 61 LOSLPTHPSPLORYSEDPYPLPSETDGYVAPLTCSPOPEYVNPDPVPPSPREGPL 120
 Db 1066 LOSLPTHPSPLORYSEDPYPLPSETDGYVAPLTCSPOPEYVNPDPVPPSPREGPL 1125
 Oy 121 PAARAGATLERAKTILSPGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180
 Db 1126 PAARAGATLERAKTILSPGKNGVYKDVAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 1165
 Oy 181 DNLTYWDDPPERGAPSTFKGTPTAENPEYLTGLDVP 217
 Db 1165 DNLTYWDDPPERGAPSTFKGTPTAENPEYLTGLDVP 1179

Db	1186	DNLVYMDODPPERCAPSTFKGPTAENPEYLGUDVP	1222
RESULT	9		
ID	AAV92620		
XX	AAV92620	standard; Protein: 1255 AA.	
XX	AAV92620;		
XX	10-AUG-2000	(first entry)	
DE	Human heregulin 2 (Her2).		
XX			
KW	Heregulin 2; Her2; vaccination; cytotoxic T-lymphocyte immunity		
KM	self-protein; cancer; breast cancer; prostate cancer;		
XX	cell-associated peptide antigen; foreign epitope.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	1..173		
FT	/label="N-terminal		
FT	/note="mature polypeptide"		
FT	5..25		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	59..73		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	103..117		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	149..163		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	174..323		
FT	/label="Cysteine_rich_domain		
FT	210..224		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	250..264		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	324..483		
FT	/label="Ligand_binding_domain		
FT	325..339		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	369..383		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	465..479		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	484..623		
FT	/label="Cysteine_rich_domain		
FT	579..593		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	624..654		
FT	/label="Transmembrane_domain		
FT	632..652		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	653..667		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	655..1010		
FT	/label="Tyrosine_kinase_domain		
FT	661..675		
FT	/label="insertion_region		
FT	/note="suitable for foreign epitope insertion"		
FT	695..709		

FT	/label= insertion_region
FT	/note= "suitable for foreign epitope insertion"
FT	710..730
FT	/label= insertion_region
FT	/note= "suitable for foreign epitope insertion"
FT	1011..1235
FT	/label= C-terminal_domain
PX	
PN	WO200020027-A2.
XX	
PD	13-APR-2000.
XX	
PF	05-OCT-1999;
XX	99WO-DK00525.
PR	05-OCT-1998;
XX	98DK-0001261.
PP	20-OCT-1998;
XX	98US-0105011.
PA	(MEBT-) M & E BIOTECH AS.
XX	
PI	Steinae L, Mouritsen S, Nielsen KG, Haaning J, Leach D, Dalum I;
P1	Gautam A, Birk P, Karlsson G;
DR	WPI: 2000-349917/30.
DR	N-PSTDB; AAA09455.
PT	Inducing immune responses to weakly immunogenic, tumor associated
PT	peptide antigens for the treatment of breast and prostate cancer
PS	Claim 62; Page 193-198; 220pp; English.
XX	
CC	This is the human heregulin 2 (Her2) sequence. Immunogenic analogues of
CC	Her2 can be used in the claimed method as an autovaccine to induce a CTL
CC	response. Subdominant CTL epitopes, antibody binding regions and
CC	cysteine residues involved in disulfide bonds are preserved in the
CC	immunogenized forms. Regions suitable for the insertion of foreign T
CC	helper epitopes were identified (see features table). The method
CC	is used for inducing immune responses against weakly immunogenic
CC	cell-associated peptide antigens (PA) such as those associated with
CC	cancers (self-proteins), e.g. human prostate specific membrane antigen
CC	(PSM), heregulin 2 (Her2) and/or fibroblast growth factor 8b (FGF8b).
CC	The method comprises effecting simultaneous presentation by antigen
CC	producing cells (APCs) of the animals immune system of: (1) at least 1
CC	CTL (cytotoxic T-lymphocyte) group derived from the PA and/or at least 1
CC	B-cell group derived from the cell-associated PA; and (2) at least 1
CC	first T helper cell group which is foreign to the animal. Analogues of
CC	human PSM, human Her2 and human/murine FGF8b comprising a substantial
CC	part of all known and predicted CTL and B-cell epitopes of the respective
CC	PA and including at least one foreign T helper epitope are also claimed.
CC	The method is used to treat prostate, prostate/breast or breast cancer
CC	when the PA is human PSM, FGF8b and Her2, respectively.
XX	
SO	Sequence 1255 AA:
Query Match	100.0%; Score 1182; DB 21; Length 1255;
Best Local Similarity	100.0%; Pred. No. 1.3e-90;
Matches 217; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GAGGAVNHRHRRSSSTRSGGDITLGLPSEEEAAPSPLAPSGAGSDVDGDLGMKAANG 60
Db	1038 GAGGVVHHRRHRSSTRSGGDITLLEPSEEAAPRPPLAPSGAGSDVDGDLGMKAANG 1097
OY	61 LOSLTTHPPSPLORTSDEPTVLPLESETDGYVAPLTCSPQPEVYNQNDVPPOPPSPREGCL 120
Db	1098 LOSLTTHPPSPLORTSDEPTVLPLESETDGYVAPLTCSPQPEVYNQNDVPPOPPSPREGCL 1157
OY	121 PAARPAGATLERAKTLSPCKNGVKVDVAFGAGAVENPEYLTPQGGAAPQHPHPAPFAFP 180
Db	1158 PAARPAGATLERAKTLSPCKNGVKVDVAFGAGAVENPEYLTPQGGAAPQHPHPAPFAFP 1217
OY	181 DNLYYWDDDPERGAAPSTFKCTPTAENPEYLGIDLP 217
Db	1218 DNLYYWDDDPERGAAPSTFKCTPTAENPEYLGIDLP 1254

```
RESULT 10
AAE12130
ID AAE12130 standard; Protein: 1255 AA.
XX
AC AAE12130:
XX
DT 18-DEC-2001 (first entry)
XX
DE Human tyrosine kinase-type receptor, HER-2.
XX
KW Therapeutic compound; major histocompatibility complex; vaccine;
KW antigenic peptide; MHC; immunoregulatory; immune response; HER-2;
KW adoptive immunotherapy; anti-cancer; breast cancer antigen; APC;
KW antigen presenting cell; human; tyrosine kinase-type receptor.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 774..782
FT /note= "Antigenic epitope"
XX
PN WO200168677-A2.
XX
PD 20-SEP-2001.
XX
PE 16-MAR-2001; 2001WO-USA0328.
XX
PR 16-MAR-2000; 2000US-0527487.
XX
PA (GEN2 ) GENZYME CORP.
XX
PI Nicotlette CA;
XX
DR WPI: 2001-616284/71.
DR N-PSDB: AAD19731.
XX
PT Novel synthetic therapeutic compound for inducing immune response and
PT for use in adoptive immunotherapy, has enhanced binding to major
PT histocompatibility molecules and enhanced immunoregulatory properties
PT
XX
PS Claim 4; Page 63-67; 69pp; English.
XX
CC The invention relates to synthetic therapeutic compounds (antigenic
CC peptides) with enhanced binding to major histocompatibility complex
CC (MHC) molecules and enhanced immunoregulatory properties relative
CC for their natural counterparts. Compounds of the invention are useful
CC for inducing an immune response in a subject and for use in adoptive
CC immunotherapy. They are useful as components of anti-cancer vaccines
CC and to expand immune effector cells that are specific for cancers
CC characterised by expression of the breast cancer antigen, HER-2.
CC Polynucleotides that encode peptides of the invention are useful as
CC hybridisation probes and as primers for the detection of genes of gene
CC transcripts that are expressed in antigen presenting cells (APCs), to
CC confirm transduction of polynucleotides into host cells. The present
CC sequence is human tyrosine kinase-type receptor, HER-2. Compounds
CC of the invention are designed based on the HER-2 antigenic peptide
CC (774-782).
XX
SQ Sequence 1255 AA;
XX
Query Match 100.0%; Score 1182; DB 22; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.3e-90;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPAPSEAGSDVFDGDLGGAAG 60
DB 1038 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPAPSEAGSDVFDGDLGGAAG 1097
QY 61 LQSLPTHPSPLQRYSEDPVPLPSETDGYAAPLTCSPQPEYVNPDPVRPQPPSPREGPL 120
DB 1098 LQSLPTHPSPLQRYSEDPVPLPSETDGYAAPLTCSPQPEYVNPDPVRPQPPSPREGPL 1157
```

```
QY 121 PAARAGATLERAKTSLSPKNGVYKDVAFGGAVERNPEYLPQGGAPQPHPPAFSPAF 180
DB 1158 PAARAGATLERAKTSLSPKNGVYKDVAFGGAVERNPEYLPQGGAPQPHPPAFSPAF 1217
QY 181 DNLVYWDODPPERGAPSPSTFGKTPTAENPEYLGIDVP 217
DB 1218 DNLVYWDODPPERGAPSPSTFGKTPTAENPEYLGIDVP 1254
RESULT 11
AAB60167
ID AAB60167 standard; Protein: 1255 AA.
XX
AC AAB60167;
XX
DT 03-APR-2001 (first entry)
XX
DE HER2 transgene plasmid construct encoded protein.
XX
KW Human; HER2; ErbB2 receptor; p185neu; maytansinoid conjugate; cancer;
KW antibody.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200100244-A2.
XX
PD 04-JAN-2001.
XX
PE 23-JUN-2000; 2000WO-US17229.
XX
PR 25-JUN-1999; 99US-0141316.
PR 16-MAR-2000; 2000US-0189844.
XX
PA (GETH ) GENENTECH INC.
XX
PI Erickson S, Schwall R;
XX
DR WPI: 2001-061962/07.
DR N-PSDB: AAF24297.
XX
PT Treating tumors, particularly breast cancers, which overexpress an ErbB
PT receptor and does not respond to an anti-ErbB antibody, comprises
PT conjugating the antibody to a maytansinoid -
XX
PS Example 3; Fig 4; 92pp; English.
XX
CC The present invention provides a method of treating cancer by
CC administering a conjugate of anti-ErbB antibody with a maytansinoid. In
CC particular, the antibody is directed against ErbB2 (also known as HER2
CC and p185neu). The method is particularly useful in the treatment of
CC breast, ovarian, stomach, endometrial, salivary gland, lung, kidney,
CC colon, colorectal, thyroid, pancreatic, prostate and bladder cancers.
XX
SQ Sequence 1255 AA;
XX
Query Match 100.0%; Score 1182; DB 22; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.3e-90;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPAPSEAGSDVFDGDLGGAAG 60
DB 1038 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPAPSEAGSDVFDGDLGGAAG 1097
QY 61 LQSLPTHPSPLQRYSEDPVPLPSETDGYAAPLTCSPQPEYVNPDPVRPQPPSPREGPL 120
DB 1098 LQSLPTHPSPLQRYSEDPVPLPSETDGYAAPLTCSPQPEYVNPDPVRPQPPSPREGPL 1157
QY 121 PAARAGATLERAKTSLSPKNGVYKDVAFGGAVERNPEYLPQGGAPQPHPPAFSPAF 180
DB 1158 PAARAGATLERAKTSLSPKNGVYKDVAFGGAVERNPEYLPQGGAPQPHPPAFSPAF 1217
```

```

0Y 181 DNLVYWDOPPERGAPPSFTEKSTPTAENEPEYGLDVP 217
   |||||||||||||||||||||||||||||||||||||||
DB 1218 DNLVYWDOPPERGAPPSFTEKSTPTAENEPEYGLDVP 1254

RESULT 12
ID AU074545
AU074545
AAU074545 standard; Protein; 1255 AA.
AC AU074545;
AD 23-APR-2002 (first entry)
DE Human HER2 (ErbB2) polypeptide.
KW Human; HER2; ErbB; epidermal growth factor receptor; receptor;
   anti-ErbB antibody-maytensinoid conjugate; cancer; tumour; breast; ovary;
   stomach; oesophagus; salivary gland; lung; kidney; colon; colorectum;
   thyroid; pancreas; prostate; bladder; ErbB2; neuronal disorder;
   glial disorder; astrocytal disorder; hypothalamic disorder;
   glandular disorder; macrophagal disorder; epithelial disorder;
   stromal disorder; blastocoealic disorder; inflammatory disorder;
   angiogenic disorder; immunological disorder.
KW Homo sapiens.
OS US2002001587-A1.
PN 03-JAN-2002.
PF 16-MAR-2001; 2001US-0811123.
PR 16-MAR-2000; 2000US-189844P.
PR 05-OCT-2000; 2000US-238327P.
XX (ERIC/) ERICKSON S.
XX (SCHW/) SCHWALL R.
XX (SLIW/) SLIWKOWSKI M.
XX
XX Erickson S, Schwall R, Sliwkowski M;
XX WPI; 2002-163686/21.
XX DR N-PSDB; ABK14058.
XX
XX Treating tumour characterised by overexpression of epidermal growth
PT factor receptor, ErbB or cancer in mammal, comprises administering
PT anti-ErbB antibody-maytensinoid conjugate to the mammal
XX
XX Example 3; Fig 7; 93pp; English.
XX
XX The invention relates to treating a tumour in a mammal, where the tumour
CC is characterised by the overexpression of an epidermal growth factor
CC receptor (ErbB) and does not respond or responds poorly, to treatment
CC with an anti-ErbB antibody, comprising administering to the mammal an
CC anti-ErbB antibody-maytensinoid conjugate. The method is useful for
CC treating cancer or tumours of the breast, ovary, stomach, endometrium,
CC salivary gland, lung, kidney, colon, colorectum, thyroid, pancreas,
CC prostate and bladder, preferably breast cancer. The breast cancer is a
CC metastatic breast cancer or an aggressive form of metastatic breast
CC cancer which overexpresses ErbB2. The method is also useful for treating
CC neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,
CC epithelial, stromal, blastocoealic, inflammatory, angiogenic and
CC immunological disorders. This sequence represents the human HER2 (ErbB2)
CC polypeptide of the invention.
CC
CC Sequence 1255 AA;
SO

Query Match 100.0%; Score 1182; DB 23; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1,3e-90;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0

1 GAGCGVHHRRSSSRSGGDLTTLGLEPSEEARPSPLAPSGAGSDVVDGLGMAANG 60
|||||||||||||||||||||||||||||||||||||

```

```

Db      1038 GAGGAWHHNRHSSSTRSGGGDLTLGLPEEERAPRSLPASEACSDVFDGDLMGAANG 1097
Oy      61   LOSLPHDPSPLQRRSEDPTVPLPSTDDGVAVPLTSCFQPEYVNODVDVPQPSPREBGL 120
Db      1098 LOSLPHDPSPLQRRSEDPVPLPSETDDYVAFLTCSPPQPEYVNODVDVRQPSPREGPL 1157
Oy      121   PARAPGATLEBAKTLSPKNGVVKVFAPFGAVENTPELTPOGGAAPQHPPPAFSPAF 180
Db      1156 PARAPGATLEBAKTLSPKNGVVKDFAPFGAVEPELTPOGGAAPQHPPPAFSPAF 1217
Oy      181 DNLYYWDODPPERGAPPSTFKGTPTAEENDEYLGLDVP 217
Db      1218 DNLYYWDODPPERGAPPSTFKGTPTAEENDEYLGLDVP 1254

RESULT 13
AAB21201
ID      AAB21201 standard; protein: 266 AA.
XX
AC      AAB21201;
XX
DT      12-JAN-2001 (first entry)
XX
DE      Human HER-2/neu protein phosphorylation domain.
XX
KM      Human; HER-2/neu; oncogene; tyrosine kinase; cytosolic; vaccine;
KM      breast cancer; prostate cancer; ovarian cancer; lung cancer;
KM      colon cancer.
XX
OS      Homo sapiens.
XX
PN      WO200044899-A1.
PD      03-AUG-2000.
PF      28-JAN-2000; 2000MO-US02164.
PR      29-JAN-1999; 99US-0117976.
PA      (CORI-) CORIXA CORP.
PA      (SMIK ) SMITHKLINE BEECHAM.
PI      Cheever MA, Gheysen D;
PI      WPI: 2000-505976/45.
DR      HER-2/neu extracellular domain/phosphorylation domain fusion proteins
PT      useful for vaccinating against breast, ovarian, colon, lung and
PT      prostate cancers -
PS      Claim 2; Fig 10; 128pp; English.
XX
CC      The present sequence is the phosphorylation domain of the HER-2/neu
CC      protein. HER-2/neu is a member of the tyrosine kinase family of
CC      receptor-like glycoproteins and shows homology to the epidermal growth
CC      factor receptor (EGFR). It probably plays a part in cell-growth and/or
CC      differentiation. The HER-2/neu gene is an oncogene. An HER-2/neu fusion
CC      protein comprising a HER-2/neu extracellular domain fused to a HER-2/neu
CC      phosphotyrlation domain may be used to treat or prevent cancer by
CC      eliciting or enhancing an immune response to the HER-2/neu protein. It
CC      may be used to treat malignancies such as breast, ovarian, colon, lung
CC      and prostate cancers, and may be used as an antigen to vaccinate against
CC      these neoplasias.
CX      xx
SQ      Sequence 266 AA;

Query Match          99.6%; Score 1177; DB 21; Length 266;
Best Local Similarity 99.5%; Pred. No. 6e-91;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

1 GAGGAWHHNRHSSSTRSGGGDLTLGLPEEERAPRSLPASEACSDVFDGDLMGAANG 60
49 GAGGAWHHNRHSSSTRSGGGDLTLGLPEEERAPRSLPASEACSDVFDGDLMGAANG 108

```

QY 61 LQSLPTHDPSLQRYSEDPVPLPSETDGYVAPLTCSQPEYVNOVDVRPQPPSPREGPL 120
 DB 109 LQSLPTHDPSLQRYSEDPVPLPSETDGYVAPLTCSQPEYVNOVDVRPQPPSPREGPL 168
 QY 121 PAARPAGATLTERAKTSLSPGKNGVYKDVAFSGAVENPEYLPQGGAAPOPHPPAFSPAF 180
 DB 169 PAARPAGATLTERAKTSLSPGKNGVYKDVAFSGAVENPEYLPQGGAAPOPHPPAFSPAF 228
 QY 181 DNLVYWDQDPPERGAPSTFGKGTPTAENPEYLGLDVP 217
 DB 229 DNLVYWDQDPPERGAPSTFGKGTPTAENPEYLGLDVP 265

RESULT 14
 ID AAM51146 standard; Protein; 266 AA.
 XX AAM51146;
 AC AAM51146;
 XX 17-JUN-2002 (first entry)
 DT 17-JUN-2002 (first entry)
 DE Human Her-2/neu oncoprotein phosphorylation domain.
 XX Her-2/neu; oncogene; cancer; tumour; vaccine; human; p185;
 KM tyrosine kinase; receptor; c-erbB2; gene therapy.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200212341-A2.
 PD 14-FEB-2002.
 PE 03-AUG-2001; 2001WO-US24283.
 PF 03-AUG-2000; 2000US-0632507.
 PR 03-AUG-2000; 2000US-0632507.
 PA (CORI-) CORIXA CORP.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Cheever MA, Gheysen D;
 PI WPI: 2002-241743/29.
 DR Her-2/neu fusion protein for treating or preventing cancer by eliciting
 PT or enhancing an immune response to the protein, has Her-2/neu
 PT extracellular domain fused to Her-2/neu intracellular or
 PT phosphorylation domain
 PS Claim 2; Fig 10; 141pp; English.

CC The present sequence is that of the phosphorylation domain of
 CC human Her-2/neu (p185 glycoprotein or c-erbB2), an oncogenic
 CC self-protein and target for anti-cancer vaccines. The Her-2/neu
 CC gene is amplified and p185 is overexpressed in a variety of cancers,
 CC including breast, ovarian, colon, lung and prostate cancer.
 CC Her-2/neu (see AAM51143) is a member of the tyrosine kinase family
 CC of receptor-like glycoproteins. It comprises an extracellular
 CC domain with homology to the epidermal growth factor receptor
 CC (EGFR), a highly hydrophobic transmembrane domain and a C-terminal
 CC intracellular domain that also shows homology to EGFR. Its
 CC overexpression correlates with a poor prognosis in breast and
 CC ovarian cancers. The invention provides Her-2/neu fusion
 CC proteins, nucleic acids encoding them, viral vectors, and vaccines
 CC comprising the fusion proteins or nucleic acid molecules. In
 CC preferred fusion proteins, the extracellular domain of a Her-2/neu
 CC protein is fused to a Her-2/neu intracellular domain or
 CC phosphorylation domain (or its DeltaPD fragment). An immune
 CC response to Her-2/neu protein is elicited or enhanced by
 CC administering the fusion protein in the form of a vaccine, or by
 CC transfecting cells of an animal ex vivo with a nucleic acid
 CC encoding the fusion protein, and delivering the transfected cells
 CC to the animal. The fusion proteins, nucleic acids, and isolated

CC specific T-cells are useful for inhibiting the development of a
 CC cancer, especially breast, ovarian, colon, lung or prostate cancer
 CC in a patient. T cells that specifically react with a Her-2/neu
 CC fusion protein can be used to remove tumour cells from a sample in
 CC order to inhibit the development of cancer in a patient.

SQ Sequence 266 AA;
 Query Match 99.6%; Score 1177; DB 23; Length 266;
 Best Local Similarity 99.5%; Pred. No. 66-91; Mismatches 1; Gaps 0;
 Matches 216; Conservative 0; Indels 1; Indels 0;

QY 1 GAGGCVHHRRSSSTRSGGDLTGLPESEEARPSPLAPSEGASDVPDGLGCAAGK 60
 DB 49 GAGGCVHHRRSSSTRSGGDLTGLPESEEARPSPLAPSEGASDVPDGLGCAAGK 108
 QY 61 LQSLPTHDPSLQRYSEDPVPLPSETDGYVAPLTCSQPEYVNOVDVRPQPPSPREGPL 120
 DB 109 LQSLPTHDPSLQRYSEDPVPLPSETDGYVAPLTCSQPEYVNOVDVRPQPPSPREGPL 168
 QY 121 PAARPAGATLTERAKTSLSPGKNGVYKDVAFSGAVENPEYLPQGGAAPOPHPPAFSPAF 180
 DB 169 PAARPAGATLTERAKTSLSPGKNGVYKDVAFSGAVENPEYLPQGGAAPOPHPPAFSPAF 228
 QY 181 DNLVYWDQDPPERGAPSTFGKGTPTAENPEYLGLDVP 217
 DB 229 DNLVYWDQDPPERGAPSTFGKGTPTAENPEYLGLDVP 265

RESULT 15
 ID AAE20483 standard; Protein; 583 AA.
 XX AAE20483;
 AC AAE20483;
 XX 01-JUL-2002 (first entry)
 DT 01-JUL-2002 (first entry)
 DE Human protein for the clone HICD_native_coding_region.
 XX Human; Her-2/Neu protein; immune response; gene therapy; breast cancer;
 KM human leukocyte antigen; HLA; vaccine; malignancy; cytostatic.
 XX Homo sapiens.
 OS Homo sapiens.
 PN Key Location/Qualifiers
 FT MISC-difference 581..583
 FT /note="Encoded by GTGTAATGATC"

MO200214503-A2.
 PD 21-FEB-2002.
 PE 14-AUG-2001; 2001WO-US41733.
 PF 14-AUG-2000; 2000US-225152P.
 PR 28-SEP-2000; 2000US-236428P.
 PR 21-FEB-2001; 2001US-270520P.
 PA (CORI-) CORIXA CORP.
 PA Hand-Zimmermann S, Cheever MA, Foy TM, Lodes MJ, Kalos ND;
 PI McNeill PD, Vedvick TS;
 PI WPI: 2002-280758/32.
 DR N-PSDB: AAD32746.
 DR Novel isolated Her-2/Neu polypeptide composition useful for therapy,
 PT prevention and diagnosis of cancer, preferably breast cancer
 PT Example 5; Page 121-122; 129pp; English.
 PS The invention relates to an isolated Her-2/Neu polypeptide composition
 CC effective for eliciting an immune response. The invention is useful for

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:40:24 ; Search time 7.98328 Seconds

(without alignments)
2178.074 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182

Sequence: 1 GAGCMVHHRRSSSTRSGCG.....STFKGPTAENPEYLGLDVP 217

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications-AA:*

1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182	100.0	217	10	US-09-821-883-25
2	1182	100.0	397	10	US-09-821-883-27
3	1182	100.0	555	10	US-09-821-883-1
4	1182	100.0	564	10	US-09-821-883-3
5	1182	100.0	680	10	US-09-821-883-2
6	1182	100.0	697	10	US-09-821-883-4
7	1182	100.0	1179	10	US-09-821-883-29
8	1182	100.0	1255	10	US-09-769-508-2
9	1182	100.0	1255	10	US-09-811-123-9
10	1182	100.0	1255	10	US-09-811-115-3
11	1177	99.6	266	9	US-09-854-356-4
12	1177	99.6	353	9	US-09-930-125-9
13	1177	99.6	587	9	US-09-930-125-8
14	1177	99.6	589	9	US-09-930-125-10
15	1177	99.6	600	9	US-09-930-125-11
16	1177	99.6	919	9	US-09-854-356-6
17	1177	99.6	1255	9	US-09-854-356-1
18	1177	99.6	1255	9	US-09-930-125-2
19	1177	99.6	1255	9	US-09-441-411-6

20	961	81.3	1256	9	US-09-854-356-14	Sequence 14, App1
21	954	80.7	1256	9	US-09-854-356-2	Sequence 2, App1
22	954	80.7	1260	9	US-09-870-759-118	Sequence 118, App
23	711	60.2	293	9	US-10-102-806-583	Sequence 583, App
24	648.5	54.9	135	10	US-09-925-301-1232	Sequence 1232, App
25	164	13.9	1210	10	US-09-725-433-2	Sequence 2, App1
26	157	13.3	1308	10	US-09-940-101-2	Sequence 4, App1
27	155	13.1	705	9	US-09-158-722-4	Sequence 2, App1
28	134	11.3	674	9	US-10-086-464-14	Sequence 14, App1
29	127.5	10.8	507	9	US-10-078-547-24	Sequence 24, App1
30	126	10.7	1274	9	US-10-020-215-2	Sequence 2, App1
31	124	10.5	503	9	US-10-078-547-2	Sequence 2, App1
32	121	10.2	428	9	US-09-906-514-4	Sequence 4, App1
33	120	10.2	699	9	US-10-121-988-143	Sequence 143, App
34	118.5	10.0	731	9	US-10-086-464-8	Sequence 8, App1
35	118	10.0	419	10	US-09-814-777A-36	Sequence 36, App1
36	118	10.0	684	10	US-09-823-240-9	Sequence 9, App1
37	115.5	9.8	731	9	US-10-086-464-17	Sequence 17, App1
38	115.5	9.8	743	10	US-09-771-161A-164	Sequence 164, App
39	115.5	9.8	743	10	US-09-771-161A-254	Sequence 254, App
40	114	9.6	1464	9	US-10-060-036-159	Sequence 159, App1
41	113.5	9.6	726	10	US-09-770-689A-4	Sequence 4, App1
42	113.5	9.6	803	10	US-09-770-689A-2	Sequence 2, App1
43	113.5	9.6	1665	10	US-09-858-664A-2	Sequence 2, App1
44	113.5	9.6	2630	9	US-10-077-130-2	Sequence 2, App1
45	113.5	9.6	7968	9	US-10-077-130-5	Sequence 5, App1

ALIGNMENTS

```

RESULT 1
US-09-821-883-25
; Sequence 25, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Thomas
; APPLICANT: Gradovic, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821, 883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-821-883-25

Query Match      100.0%; Score 1182; DB 10; Length 217;
Best Local Similarity 100.0%; Pred. No. 5.3e-70;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 2

US-09-821-883-27

: Sequence 27, Application US/09821883
: Patent No. US20020061310A1

: GENERAL INFORMATION:

: APPLICANT: Laus, Reiner

: APPLICANT: Vidovic, Damir

: APPLICANT: Gradis, Thomas

: TITLE OF INVENTION: Compositions and Methods for Dendritic

: FILE REFERENCE: 7636-0022.30

: CURRENT APPLICATION NUMBER: US/09/821,883

: PRIOR FILING DATE: 2001-03-30

: PRIOR FILING DATE: 2000-03-30

: NUMBER OF SEQ ID NOS: 30

: SOFTWARE: FastSeq for Windows Version 4.0

: SEQ ID NO 27

: LENGTH: 397

: TYPE: PRT

: ORGANISM: Artificial Sequence

: FEATURE:

: OTHER INFORMATION: NY-ESO-1C tumor antigen

US-09-821-883-27

Query Match 100.0%; Score 1182; DB 10; Length 397;
Best Local Similarity 100.0%; Pred. No. 1e-69;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

: DB 1 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAAG 60

: DB 181 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAAG 240

: DB 61 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAALTCSPQEVYNQDPVRQPPSPREGPL 120

: DB 241 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAALTCSPQEVYNQDPVRQPPSPREGPL 300

: DB 121 PAARPGATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180

: DB 301 PAARPGATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 360

: DB 181 DNLTYWDQDPPERGAAPSTFKGTPTAENPEYLGIDVP 217

: DB 361 DNLTYWDQDPPERGAAPSTFKGTPTAENPEYLGIDVP 397

RESULT 3

US-09-821-883-1

: Sequence 1, Application US/09821883
: Patent No. US20020061310A1

: GENERAL INFORMATION:

: APPLICANT: Laus, Reiner

: APPLICANT: Vidovic, Damir

: APPLICANT: Gradis, Thomas

: TITLE OF INVENTION: Compositions and Methods for Dendritic

: FILE REFERENCE: 7636-0022.30

: CURRENT APPLICATION NUMBER: US/09/821,883

: PRIOR FILING DATE: 2001-03-30

: PRIOR FILING DATE: 2000-03-30

: NUMBER OF SEQ ID NOS: 30

: SOFTWARE: FastSeq for Windows Version 4.0

: SEQ ID NO 1

: LENGTH: 555

: TYPE: PRT

: ORGANISM: Artificial Sequence

: FEATURE:

: OTHER INFORMATION: HER500 construct

US-09-821-883-1

Query Match 100.0%; Score 1182; DB 10; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.4e-69;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

: DB 1 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAAG 60

: DB 330 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAAG 389

: DB 61 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAALTCSPQEVYNQDPVRQPPSPREGPL 120

: DB 390 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAALTCSPQEVYNQDPVRQPPSPREGPL 449

: DB 121 PAARPGATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180

: DB 450 PAARPGATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 509

: DB 181 DNLTYWDQDPPERGAAPSTFKGTPTAENPEYLGIDVP 217

: DB 510 DNLTYWDQDPPERGAAPSTFKGTPTAENPEYLGIDVP 546

RESULT 4

US-09-821-883-3

: Sequence 3, Application US/09821883
: Patent No. US20020061310A1

: GENERAL INFORMATION:

: APPLICANT: Laus, Reiner

: APPLICANT: Vidovic, Damir

: APPLICANT: Gradis, Thomas

: TITLE OF INVENTION: Compositions and Methods for Dendritic

: FILE REFERENCE: 7636-0022.30

: CURRENT APPLICATION NUMBER: US/09/821,883

: PRIOR FILING DATE: 2001-03-30

: PRIOR FILING DATE: 2000-03-30

: NUMBER OF SEQ ID NOS: 30

: SOFTWARE: FastSeq for Windows Version 4.0

: SEQ ID NO 3

: LENGTH: 564

: TYPE: PRT

: ORGANISM: Artificial Sequence

: FEATURE:

: OTHER INFORMATION: HER500+ construct

US-09-821-883-3

Query Match 100.0%; Score 1182; DB 10; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.4e-69;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

: DB 1 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAAG 60

: DB 339 GAGGWHHRSSSTRSGGDLTLGLEPSEEARPSRLAPSEAGSDVFDGDLGMAAG 398

: DB 61 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAALTCSPQEVYNQDPVRQPPSPREGPL 120

: DB 399 LQSLPTHDPSPDLQRYSDPVPPLPSETDGYVAALTCSPQEVYNQDPVRQPPSPREGPL 458

: DB 121 PAARPGATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 180

: DB 459 PAARPGATLERAKTLPSPKNGVYKDVFAFGAVENPEYLTPOGGAAPQHPPPAFSPAF 518

: DB 181 DNLTYWDQDPPERGAAPSTFKGTPTAENPEYLGIDVP 217

: DB 519 DNLTYWDQDPPERGAAPSTFKGTPTAENPEYLGIDVP 555

RESULT 5

US-09-821-883-2

: Sequence 2, Application US/09821883
: Patent No. US20020061310A1

: GENERAL INFORMATION:

: APPLICANT: Laus, Reiner

US-09-821-883-2

APPLICANT: Vidovic, Damir
APPLICANT: Gradidis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE OF INVENTION: Cell-Based Immunotherapy
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 690
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500-hgm-CSF construct
US-09-821-883-2

Query Match 100.0%; Score 1182; DB 10; Length 690;
Best Local Similarity 100.0%; Pred. No. 1,8e-69;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEGAGSDVFDGDLGKAAG 60
DB 330 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEGAGSDVFDGDLGKAAG 389
QY 61 LQSLPTHPSPLQRYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 120
DB 390 LQSLPTHPSPLQRYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 449
QY 121 PAAPGATLERAKTILSGKNGVYKDVAFGAVENPEYLPQGGAPQPHPPAFSPAF 180
DB 450 PAAPGATLERAKTILSGKNGVYKDVAFGAVENPEYLPQGGAPQPHPPAFSPAF 509
QY 181 DNLVYWDODPPERGAPSTFGKPTAENPEYLGLDVP 217
DB 510 DNLVYWDODPPERGAPSTFGKPTAENPEYLGLDVP 546

RESULT 6
US-09-821-883-4
Sequence 4, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Relner
APPLICANT: Vidovic, Damir
APPLICANT: Gradidis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 697
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: HER500-rgm-CSF construct
US-09-821-883-4

Query Match 100.0%; Score 1182; DB 10; Length 697;
Best Local Similarity 100.0%; Pred. No. 1,8e-69;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEGAGSDVFDGDLGKAAG 60
DB 330 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEGAGSDVFDGDLGKAAG 398

QY 61 LQSLPTHPSPLQRYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 120
DB 399 LQSLPTHPSPLQRYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 458
QY 121 PAAPGATLERAKTILSGKNGVYKDVAFGAVENPEYLPQGGAPQPHPPAFSPAF 180
DB 459 PAAPGATLERAKTILSGKNGVYKDVAFGAVENPEYLPQGGAPQPHPPAFSPAF 518
QY 181 DNLVYWDODPPERGAPSTFGKPTAENPEYLGLDVP 217
DB 519 DNLVYWDODPPERGAPSTFGKPTAENPEYLGLDVP 555

RESULT 7
US-09-821-883-29
Sequence 29, Application US/09821883
Patent No. US20020061310A1
GENERAL INFORMATION:
APPLICANT: Laus, Relner
APPLICANT: Vidovic, Damir
APPLICANT: Gradidis, Thomas
TITLE OF INVENTION: Compositions and Methods for Dendritic
FILE REFERENCE: 7636-0022.30
CURRENT APPLICATION NUMBER: US/09/821,883
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193,504
PRIOR FILING DATE: 2000-03-30
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29
LENGTH: 1179
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: SART-3-IC
US-09-821-883-29

Query Match 100.0%; Score 1182; DB 10; Length 1179;
Best Local Similarity 100.0%; Pred. No. 3,1e-69;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEGAGSDVFDGDLGKAAG 60
DB 963 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLAPSEGAGSDVFDGDLGKAAG 1022
QY 61 LQSLPTHPSPLQRYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 120
DB 1023 LQSLPTHPSPLQRYSEDPYPLPSETDGYAAPLTCSPQPEYVNPDPVPPSPREGPL 1082
QY 121 PAAPGATLERAKTILSGKNGVYKDVAFGAVENPEYLPQGGAPQPHPPAFSPAF 180
DB 1083 PAAPGATLERAKTILSGKNGVYKDVAFGAVENPEYLPQGGAPQPHPPAFSPAF 1142
QY 181 DNLVYWDODPPERGAPSTFGKPTAENPEYLGLDVP 217
DB 1143 DNLVYWDODPPERGAPSTFGKPTAENPEYLGLDVP 1179

RESULT 8
US-09-769-508-2
Sequence 2, Application US/09769508
Patent No. US20020155527A1
GENERAL INFORMATION:
APPLICANT: STUART, SUSAN G.
APPLICANT: MONAHAN, JOHN J.
APPLICANT: LANGTON, BEATRICE CLAUDIA
APPLICANT: HANCOCK, MIRIAM E.C.
APPLICANT: CHAO, LORRINE A.
APPLICANT: BLUFORD, PETER
TITLE OF INVENTION: C-ERBB-2 EXTERNAL DOMAIN: GP75
FILE REFERENCE: BB10-111-C1
CURRENT APPLICATION NUMBER: US/09/769,508

Query Match 99.6%; Score 1177; DB 9; Length 266;
Best Local Similarity 99.5%; Pred. No. 1.4e-69;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 60
|||||
DB 49 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 108
|||||

OY 61 LOSLTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVROPSPREGPL 120
|||||
DB 109 LOSLTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVROPSPREGPL 168
|||||

OY 121 PAARAGATLERAKTLSGKNGVVDVFAFGAVENPEYLPQGAAPQHPHPPAFSPAF 180
|||||
DB 169 PAARAGATLERAKTLSGKNGVVDVFAFGAVENPEYLPQGAAPQHPHPPAFSPAF 228
|||||

OY 181 DNLVYWDODPPERGAPSPSTFKGTPTAENPEYLGLDVP 217
|||||
DB 229 DNLVYWDODPPERGAPSPSTFKGTPTAENPEYLGLDVP 265
|||||

RESULT 12

US-09-930-125-9
Sequence 9, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vegdick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930.125
CURRENT FILING DATE: 2001-08-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 9
LENGTH: 583
TYPE: PRT
ORGANISM: Homo sapiens
US-09-930-125-9

Query Match 99.6%; Score 1177; DB 9; Length 583;
Best Local Similarity 99.5%; Pred. No. 3.1e-69;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 60
|||||
DB 364 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 423
|||||

OY 61 LOSLTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVROPSPREGPL 120
|||||
DB 424 LOSLTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVROPSPREGPL 483
|||||

OY 121 PAARAGATLERAKTLSGKNGVVDVFAFGAVENPEYLPQGAAPQHPHPPAFSPAF 180
|||||
DB 484 PAARAGATLERAKTLSGKNGVVDVFAFGAVENPEYLPQGAAPQHPHPPAFSPAF 543
|||||

OY 181 DNLVYWDODPPERGAPSPSTFKGTPTAENPEYLGLDVP 217
|||||
DB 544 DNLVYWDODPPERGAPSPSTFKGTPTAENPEYLGLDVP 580
|||||

RESULT 13
US-09-930-125-8
Sequence 8, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan

APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vegdick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930.125
CURRENT FILING DATE: 2001-08-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 8
LENGTH: 587
TYPE: PRT
ORGANISM: Homo sapiens
US-09-930-125-8

Query Match 99.6%; Score 1177; DB 9; Length 587;
Best Local Similarity 99.5%; Pred. No. 3.2e-69;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 60
|||||
DB 364 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 423
|||||

OY 61 LOSLTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVROPSPREGPL 120
|||||
DB 424 LOSLTHPSPLOKRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVROPSPREGPL 483
|||||

OY 121 PAARAGATLERAKTLSGKNGVVDVFAFGAVENPEYLPQGAAPQHPHPPAFSPAF 180
|||||
DB 484 PAARAGATLERAKTLSGKNGVVDVFAFGAVENPEYLPQGAAPQHPHPPAFSPAF 543
|||||

OY 181 DNLVYWDODPPERGAPSPSTFKGTPTAENPEYLGLDVP 217
|||||
DB 544 DNLVYWDODPPERGAPSPSTFKGTPTAENPEYLGLDVP 580
|||||

RESULT 14

US-09-930-125-10
Sequence 10, Application US/09930125
Publication No. US20020193329A1
GENERAL INFORMATION:
APPLICANT: Hand-Zimmerman, Susan
APPLICANT: Cheever, Martin A.
APPLICANT: Foy, Teresa M.
APPLICANT: Lodes, Michael J.
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Vegdick, Thomas S.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
FILE REFERENCE: 210121.544
CURRENT APPLICATION NUMBER: US/09/930.125
CURRENT FILING DATE: 2001-08-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 10
LENGTH: 589
TYPE: PRT
ORGANISM: Homo sapiens
US-09-930-125-10

Query Match 99.6%; Score 1177; DB 9; Length 589;
Best Local Similarity 99.5%; Pred. No. 3.2e-69;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 60
|||||
DB 372 GAGGWHHRHRSSTRSGGDLTLGLEPSEEPAPSPPLAPSGAGSDVFDGDLGGAAG 431
|||||

QY	61	LOSLPTHPSPLOQYSEDPVPLPSESDGVAAPLTCSPOEYVNOQPVRRQPSPREGL	120
Db	432	LOSLPTHPSPLOQYSEDPVPLPSESDGVAAPLTCSPOEYVNOQPVRRQPSPREGL	491
QY	121	PAARPAGATLERAKTILSPGKNQVYKDVFAFGAENEBEYTLPOGGAAPOPHPAPSPAF	180
Db	492	PAARPAGATLERAKTILSPGKNQVYKDVFAFGAENEBEYTLPOGGAAPOPHPAPSPAF	551
QY	181	DNLVYWMDDPPERGAPESTFGPTAENPEYLGADV	217
Db	552	DNLVYWMDDPPERGAPESTFGPTAENPEYLGADV	568

```

RESULT 15
US-09-930-125-11
: Sequence 11. Application US/09930125
: Publication NO. US20020193329A1
: GENERAL INFORMATION:
: APPLICANT: Hand-Zimmerman, Susan
: APPLICANT: Cheever, Martin A.
: APPLICANT: Foy, Teresa M.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Kalos, Michael D.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Veevlck, Thomas S.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS
: OF HER-2/NEU-ASSOCIATED MALIGNANCIES
: FILE REFERENCE: 210121.544
: CURRENT APPLICATION NUMBER: US/09/930.125
: CURRENT FILING DATE: 2001-08-14
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 600
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-930-125-11

```

	Query Match	99.6%	Score 1177	DA 9	Length 600
	Best Local Similarity	99.5%	Pred. No. 3,2e-69		
	Matches 216	Conservative	0	Mismatches 1	Indels 0
Qy	1	GAGGVVHHRRHSSSTRSGGGDLTLCLEPSEEBAPRSPPLASGAGSDVDFDGLGMAAKG	60		
Db	383	GAGGVVHHRRHSSSTRSGGGDLTLCLEPSEEBAPRSPPLASGAGSDVDFDGLGMAAKG	442		
Qy	61	LOSLETHPSPLOARSEPTVPLPSETGGVYAPLCSPOPEYVNOPDVAPPPSPREGRL	120		
Db	443	LOSLETHPSPLOARSEPTVPLPSETGGVYAPLCSPOPEYVNOPDVAPPPSPREGRL	502		
Qy	121	PAARPAAGATLEBAKTLSPKNGVAVKDYAFAGGAVENPEYLPOGGGAAPDPPPPAPSPAF	180		
Db	503	PAARPAAGATLEBAKTLSPKNGVAVKDYAFAGGAVENPEYLPOGGGAAPDPPPPAPSPAF	562		
Qy	181	DNLVYVODDPPRGAAPSPFTKGTPTAENPEYGLDVP	217		
Db	563	DNLVYVODDPPRGAAPSPFTKGTPTAENPEYGLDVP	599		

Search completed: April 28, 2003, 13:44:07
Job time : 7.98328 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:39:54 ; Search time 6.3503 Seconds
(without alignments)
1005.423 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182

Sequence: 1 GAGCWHHRSSSTRSGGCG.....STFKGPTAENPEYLGIDVP 217

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Issued Patents-AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUOS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1177	99.6	580	1 US-08-414-417B-69	Sequence 69, Appl
2	1177	99.6	580	2 US-08-486-348A-69	Sequence 69, Appl
3	1177	99.6	580	3 US-08-468-545B-69	Sequence 69, Appl
4	1177	99.6	580	4 US-08-466-680B-69	Sequence 68, Appl
5	1177	99.6	1255	1 US-08-467-083-68	Sequence 68, Appl
6	1177	99.6	1255	1 US-08-414-417B-68	Sequence 68, Appl
7	1177	99.6	1255	2 US-08-486-348A-68	Sequence 68, Appl
8	1177	99.6	1255	2 US-08-625-101-2	Sequence 2, Appl
9	1177	99.6	1255	2 US-08-468-545B-68	Sequence 68, Appl
10	1177	99.6	1255	2 US-08-356-766-2	Sequence 2, Appl
11	1177	99.6	1255	3 US-08-466-680B-68	Sequence 68, Appl
12	1157	97.9	1255	3 US-08-484-438-8	Sequence 8, Appl
13	164	13.9	1210	2 US-08-484-438-7	Sequence 7, Appl
14	164	13.9	1210	2 US-08-475-035-4	Sequence 4, Appl
15	157	13.3	541	2 US-08-484-438-6	Sequence 6, Appl
16	157	13.3	1308	2 US-08-484-438-2	Sequence 2, Appl
17	155	13.1	705	2 US-08-456-647B-4	Sequence 4, Appl
18	155	13.1	705	2 US-08-237-401A-4	Sequence 4, Appl
19	126.5	10.7	1321	2 US-08-317-310A-64	Sequence 64, Appl
20	126	10.7	1274	4 US-09-095-443-2	Sequence 2, Appl
21	125.5	10.6	314	2 US-08-525-742-6	Sequence 6, Appl
22	116	9.8	1248	2 US-09-080-897-2	Sequence 2, Appl
23	116	9.8	1248	4 US-09-323-735-2	Sequence 2, Appl
24	115	9.7	124	4 US-08-925-237-2	Sequence 2, Appl
25	114	9.6	1057	3 US-08-931-820-1	Sequence 1, Appl
26	114	9.6	1341	3 US-08-963-825-18	Sequence 18, Appl
27	114	9.6	1341	4 US-09-500-811-18	Sequence 18, Appl

28	114	9.6	1341	4 US-09-570-573-18	Sequence 18, Appl
29	114	9.6	1341	4 US-09-548-608-18	Sequence 18, Appl
30	114	9.6	1461	4 US-09-585-887-9	Sequence 9, Appl
31	114	9.6	1461	4 US-09-289-578-9	Sequence 9, Appl
32	113.5	9.6	2972	4 US-09-579-181-2	Sequence 2, Appl
33	113.5	9.6	3118	4 US-09-579-181-1	Sequence 1, Appl
34	111	9.4	627	4 US-08-487-596-6	Sequence 6, Appl
35	111	9.4	696	3 US-08-906-865-4	Sequence 4, Appl
36	111	9.4	696	4 US-09-129-668-4	Sequence 4, Appl
37	110	9.3	1185	4 US-09-041-886-23	Sequence 23, Appl
38	110	9.3	1315	3 US-08-899-595-3	Sequence 3, Appl
39	110	9.3	1317	3 US-09-083-521-7	Sequence 7, Appl
40	109.5	9.3	1706	2 US-08-459-568-2	Sequence 2, Appl
41	109.5	9.3	1706	2 US-08-399-411-2	Sequence 2, Appl
42	109.5	9.3	1706	3 US-08-516-859A-2	Sequence 2, Appl
43	109.5	9.3	1706	4 US-09-586-472-2	Sequence 2, Appl
44	109.5	9.3	1706	4 US-09-528-706-2	Sequence 2, Appl
45	108.5	9.2	214	1 US-08-217-327-4	Sequence 4, Appl

ALIGNMENTS

```
RESULT 1
US-08-414-417B-69
; Sequence 69, Application US/08414417B
; Patent No. 5801005
;
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/414,417B
; FILING DATE: 31-MAR-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; TELEPHONE/DOCKET NUMBER: 920010.448C2
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
; US-08-414-417B-69
;
; Query Match 99.6%; Score 1177; DB 1; Length 580;
; Best Local Similarity 99.5%; Pred. No. 3.3e-96;
; Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GAGCWHHRSSSTRSGGDLTLGLEPSEEPAPSPPLASGAGSDVDPGLJMGAMG 60
DB 363 GAGCWHHRSSSTRSGGDLTLGLEPSEEPAPSPPLASGAGSDVDPGLJMGAMG 422
```

QY 61 LOSLPTHDPSPLORXSDPTVPLPSETDGVAPLTCSPQPEYVQPDVRRPQPSPREGCL 120
DB 423 LOSLPTHDPSPLORXSDPTVPLPSETDGVAPLTCSPQPEYVQPDVRRPQPSPREGCL 482
QY 121 PAARPGATLEBRKTLSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180
DB 483 PAARPGATLEBRKTLSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 542
QY 181 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217
DB 543 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 579

RESULT 2
US-08-486-348A-69
; Sequence 69, Application US/08486348A
; Patent No. 5846538
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,348A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-486-348A-69

Query Match 99.6%; Score 1177; DB 2; Length 580;
Best Local Similarity 99.5%; Pred. No. 3.3e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 543 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 579

RESULT 3
US-08-468-545B-69
; Sequence 69, Application US/08468545B
; Patent No. 5876712
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.
; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,545B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.448C5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 580 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-468-545B-69

Query Match 99.6%; Score 1177; DB 2; Length 580;
Best Local Similarity 99.5%; Pred. No. 3.3e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGMVHRRHRSSTRSGGDLTLGLEPSEEARSPRLAPSEAGSDVFPDGLMGAKG 60
DB 363 GAGGMVHRRHRSSTRSGGDLTLGLEPSEEARSPRLAPSEAGSDVFPDGLMGAKG 422
QY 61 LOSLPTHDPSPLORXSDPTVPLPSETDGVAPLTCSPQPEYVQPDVRRPQPSPREGCL 120
DB 423 LOSLPTHDPSPLORXSDPTVPLPSETDGVAPLTCSPQPEYVQPDVRRPQPSPREGCL 482
QY 121 PAARPGATLEBRKTLSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 180
DB 483 PAARPGATLEBRKTLSPKNGVYKDVAFAGAVENPEYLTPOGGAAPQHPPPAFSPAF 542
QY 181 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 217
DB 543 DNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGIDVP 579

RESULT 4
US-08-466-680B-69
; Sequence 69, Application US/08466680B
; Patent No. 6075122
; GENERAL INFORMATION:
; APPLICANT: Cheever, Martin A.
; APPLICANT: Disis, Mary L.

;; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;; TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Seed and Berry LLP
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: US
;; ZIP: 98104-7092
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/466,680B
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sharkey, Richard G.
;; REGISTRATION NUMBER: 32,629
;; REFERENCE/DOCKET NUMBER: 920010.448C4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;;
;; INFORMATION FOR SEQ ID NO: 69:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 580 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;;
US-08-466-680B-69

Query Match 99.6%; Score 1177; DB 3; Length 580;
Best Local Similarity 99.5%; Pred. No. 3,3e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPPLAPSGAGSDVFDGCGAKG 60
DB 363 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPPLAPSGAGSDVFDGCGAKG 422
OY 61 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 120
DB 423 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 482
OY 121 PAAPAGATLERAKTLSGKNGVAVDFAFGAVENPEYLPQGGAPQPHPPAFSPA 180
DB 483 PAAPAGATLERAKTLSGKNGVAVDFAFGAVENPEYLPQGGAPQPHPPAFSPA 542
OY 181 DNLVYWDODPPERGAPSTFGTPTAENPEYLGIDVP 217
DB 543 DNLVYWDODPPERGAPSTFGTPTAENPEYLGIDVP 579

RESULT 5

US-08-467-083-68
;; Sequence 68, Application US/08467083
;; Patent No. 5726023
;; GENERAL INFORMATION:
;; APPLICANT: Cheever, Martin A.
;; APPLICANT: Disis, Mary L.
;; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;; NUMBER OF SEQUENCES: 68
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Seed and Berry
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: US

;; ZIP: 98104-7092
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/467,083
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 424
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/414,417
;; FILING DATE: 06-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sharkey, Richard G.
;; REGISTRATION NUMBER: 32,629
;; REFERENCE/DOCKET NUMBER: 920010.448C2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (206) 622-4900
;; TELEFAX: (206) 682-6031
;; TELEX: 3723836 SEEDANBERY
;;
;; INFORMATION FOR SEQ ID NO: 68:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1255 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;;
US-08-467-083-68

Query Match 99.6%; Score 1177; DB 1; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8,5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPPLAPSGAGSDVFDGCGAKG 60
DB 1038 GAGGAVHHRHSSSTRSGGDLTGLPSEEEAPRSPPLAPSGAGSDVFDGCGAKG 1097
OY 61 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 120
DB 1098 LQSLPTHPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVPRQPPSPREGPL 1157
OY 121 PAAPAGATLERAKTLSGKNGVAVDFAFGAVENPEYLPQGGAPQPHPPAFSPA 180
DB 1158 PAAPAGATLERAKTLSGKNGVAVDFAFGAVENPEYLPQGGAPQPHPPAFSPA 1217
OY 181 DNLVYWDODPPERGAPSTFGTPTAENPEYLGIDVP 217
DB 1218 DNLVYWDODPPERGAPSTFGTPTAENPEYLGIDVP 1254

RESULT 6

US-08-414-417B-68
;; Sequence 68, Application US/08414417B
;; Patent No. 3801005
;; GENERAL INFORMATION:
;; APPLICANT: Cheever, Martin A.
;; APPLICANT: Disis, Mary L.
;; TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;; TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESSES:
;; ADDRESSEE: Seed and Berry LLP
;; STREET: 6300 Columbia Center, 701 Fifth Avenue
;; CITY: Seattle
;; STATE: Washington
;; COUNTRY: US
;; ZIP: 98104-7092
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:

```

: APPLICATION NUMBER: US/08/414,417B
: FILING DATE: 31-MAR-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-4900
: TELEFAX: (206) 622-4900
: INFORMATION FOR SEQ ID NO: 68:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
US-08-414-417B-68

Query Match          99.6%; Score 1177; DB 1; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8.5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVDFDGLGMGAAG 60
DB 1038 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVDFDGLGMGAAG 1097
QY 61 LOSLPFHDPSPDLQRYSEDPVPLPSETDGVAVPLTCSPOPEYVNPDPVRPQPPSPREGPL 120
DB 1098 LOSLPFHDPSPDLQRYSEDPVPLPSETDGVAVPLTCSPOPEYVNPDPVRPQPPSPREGPL 1157
QY 121 PAAPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 180
DB 1158 PAAPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 1217
QY 181 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217
DB 1218 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 7
US-08-486-348A-68
: Sequence 68, Application US/08486348A
: Patent No. 5846538
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Disis, Mary L.
: TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN
: TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
: TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED
: NUMBER OF SEQUENCES: 69
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seed and Berry LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/486,348A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 682-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 68:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
:
US-08-486-348A-68

Query Match          99.6%; Score 1177; DB 2; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8.5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVDFDGLGMGAAG 60
DB 1038 GAGGVNHHRRSSSTRSGGDLTLGLEPSEEARPSPLASEGASDVDFDGLGMGAAG 1097
QY 61 LOSLPFHDPSPDLQRYSEDPVPLPSETDGVAVPLTCSPOPEYVNPDPVRPQPPSPREGPL 120
DB 1098 LOSLPFHDPSPDLQRYSEDPVPLPSETDGVAVPLTCSPOPEYVNPDPVRPQPPSPREGPL 1157
QY 121 PAAPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 180
DB 1158 PAAPAGATLERAKTLPSPKNGVYKDVFAFGAVENPEYITPGGAAPQPHPPAPSPAR 1217
QY 181 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 217
DB 1218 DNLVYWDQPPERGAPPSTFKGTPTAENPEYLGIDVP 1254

RESULT 8
US-08-625-101-2
: Sequence 2, Application US/08625101
: Patent No. 5869445
: GENERAL INFORMATION:
: APPLICANT: Cheever, Martin A.
: APPLICANT: Disis, Mary L.
: TITLE OF INVENTION: COMPOUNDS FOR ELICITING OR ENHANCING IMMUNE
: TITLE OF INVENTION: REACTIVITY TO HER-2/neu PROTEIN FOR PREVENTION
: TITLE OF INVENTION: OR TREATMENT OF MALIGNANCIES IN WHICH THE HER-2/neu
: TITLE OF INVENTION: ONCOGENE IS ASSOCIATED
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED AND BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/625,101
: FILING DATE: 01-APR-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Sharkey, Richard G.
: REGISTRATION NUMBER: 32,629
: REFERENCE/DOCKET NUMBER: 920010.448C7
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1255 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-08-625-101-2

Query Match          99.6%; Score 1177; DB 2; Length 1255;
Best Local Similarity 99.5%; Pred. No. 8.5e-96;
Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY	1	GAGGMYHHNRHRSSTRRGGGDLTLGLPESEEPKRRPLAPSPGASGVYDDGILGMAKG	60
Db	1038	GAGGMYHHNRHRSSTRRGGGDLTLGLPESEEPKRRPLAPSEAGSDVFDGILGMAKG	1094
QY	61	LQSLPTHDPSLQRYSEDPVPLPSEMDGVVALTCSQPOEYVNDPVARPPSPREGPL	120
Db	1098	LQSLPTHDPSLQRYSEDPVPLPSEMDGVVALTCSQPEYVNDPVARPPSPREGPL	1154
QY	121	PAAPRACATLEBAKTLSLPGKNGVYKGVFAFGAVENPEYLTLPQGGAAPOPHPPAFSPAFL	180
Db	1158	PAAPRACATLEBAKTLSLPGKNGVYKGVFAFGAVENPEYLTLPQGGAAPOPHPPAFSPAFL	1214
QY	181	DNLVYMDQDPPERCAPSTFKGTPTAENPEYGLADNV	217
Db	1218	DNLVYMDQDPPERCAPSTFKGTPTAENPEYGLADNV	1254

RESULT 11

US-08-466-680B-68

Sequence 68, Application US/08466680B

Patent No. 6075122

GENERAL INFORMATION:

APPLICANT: Cheever, Martin A.

APPLICANT: Dals, Mary L.

TITLE OF INVENTION: IMMUNE REACTIVITY TO HER-2/neu PROTEIN

TITLE OF INVENTION: FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE

TITLE OF INVENTION: HER-2/neu ONCOGENE IS ASSOCIATED

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,680B

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 920010.448C4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ. ID NO: 68:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-466-680B-68

Query Match 99.6%; Score 1177; DB 3; Length 1255;
 Best Local Similarity 99.5%; Pred. No. 8.5e-96;
 Matches 216; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGGMVHRRSSSTRSGGDLTLGLEPSEEEAPRSPLASEGASDPFDGLGMAKG 60
 |||||||

DB 1038 GAGGMVHRRSSSTRSGGDLTLGLEPSEEEAPRSPLASEGASDPFDGLGMAKG 1097
 |||||||

QY 61 LQSLPTHDPSPDLQRYSEDPTVPLPSETDGYVAAPLTCSPOPEYVNOVDVRRQPPSPREGPL 120
 |||||||

DB 1098 LQSLPTHDPSPDLQRYSEDPTVPLPSETDGYVAAPLTCSPOPEYVNOVDVRRQPPSPREGPL 1157
 |||||||

QY 121 PAARPGATLERAKTLSPKNGVYKDFAFGAGVAVENPEYLTPOGGAAPQHPAPSPAP 180
 |||||||

DB 1158 PAARPGATLERAKTLSPKNGVYKDFAFGAGVAVENPEYLTPOGGAAPQHPAPSPAP 1217
 |||||||

QY 181 DNLXYMDQDPERGAPSPSTFKGTPTAENPEYGLDVP 217
 |||||||

DB 1218 DNLXYMDQDPERGAPSPSTFKGTPTAENPEYGLDVP 1254
 |||||||

RESULT 12

US-08-484-438-8

Sequence 8, Application US/08484438

Patent No. 5811098

Patent No. 5811098 5780031

GENERAL INFORMATION:

APPLICANT: PLOWMAN, Gregory D.

APPLICANT: CULOUSCOU, Jean-Michel

APPLICANT: SHOYAB, Mohammed

APPLICANT: STEGALL, Clay B.

APPLICANT: Hellstr m, Ingegerd
 APPLICANT: Hellstr m, Karl E.
 TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,438

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,442

FILING DATE: 14-OCT-1994

APPLICATION NUMBER: US 08/150,704

FILING DATE: 10-NOV-1993

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/981,165

FILING DATE: 24-NOV-1992

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: MISTOCK, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 5624-230

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-8090

TELEFAX: (212) 869-8864/9741

INFORMATION FOR SEQ. ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1255 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-484-438-8

Query Match 97.9%; Score 1157; DB 2; Length 1255;
 Best Local Similarity 99.1%; Pred. No. 5e-94;
 Matches 216; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 GAGGMVHRRSSSTRSGGDLTLGLEPSEEEAPRSPLASEGASDPFDGLGMAKG 60
 |||||||

DB 1038 GAGGMVHRRSSSTRSGGDLTLGLEPSEEEAPRSPLASEGASDPFDGLGMAKG 1097
 |||||||

QY 61 LQSLPTHDPSPDLQRYSEDPTVPLPSETDGYVAAPLTCSPOPEYVNOVDVRRQPPSPREGPL 120
 |||||||

DB 1098 LQSLPTHDPSPDLQRYSEDPTVPLPSETDGYVAAPLTCSPOPEYVNOVDVRRQPPSPREGPL 1157
 |||||||

QY 121 PAARPGATLERAKTLSPKNGVYKDFAFGAGVAVENPEYLTPOGGAAPQHPAPSPAP 180
 |||||||

DB 1158 PAARPGATLERAKTLSPKNGVYKDFAFGAGVAVENPEYLTPOGGAAPQHPAPSPAP 1217
 |||||||

QY 181 DNLXYMDQDPERGAPSPSTFKGTPTAENPEYGLDVP 217
 |||||||

DB 1218 DNLXYMDQDPERGAPSPSTFKGTPTAENPEYGLDVP 1254
 |||||||

RESULT 13

US-08-484-438-7

Sequence 7, Application US/08484438

Patent No. 5811098

Patent No. 5811098 5780031

APPLICANT: Hellstr m, Ingegerd
APPLICANT: Hellstr m, Karl E.
TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,438
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,442
FILING DATE: 14-OCT-1994
APPLICATION NUMBER: US 08/150,704
FILING DATE: 10-NOV-1993
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/981,165
FILING DATE: 24-NOV-1992
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 5624-230
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 541 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-438-6

Query Match 13.3%; Score 157; DB 2; Length 541;
Best local Similarity 24.2%; Pred. No. 3.9e-06;
Matches 62; Conservative 23; Mismatches 75; Indels 96; Gaps 10;

QY 16 RSGGGDTTGLE-PSEERAPRSPLAP-SEGAGSDVFDGDLGMAKGLQSLPTHDPSPLO 73
DB 300 RDGFGFAEGVSVYRAPRTSTIEAPVAGATAEIFDSCNGTLRKPVAPHVQEDSSTQ 359
QY 74 RYSEDPVPLPS-----ETDGYVAPLTCSPQPEYVNOQDVNRQPPSPREGPLPAARPA 126
DB 360 RYADPVTVEAPERSPGELDEEGYMTPMRDKPQOXYLNVE-----ENPFVSR-- 408
QY 127 GATLERAKTLPSPKNGVGVKDFAFGAVENPEYLTPOGGAAPQHPPPA----- 175
DB 409 -----KNGDLQ-----ALDNPETHNANSNG-----PPKADEYVNEPLYL 442
QY 176 -----FSPAFDNLVYWDDPPPERGA--PPSTFKGTPT----- 205
DB 443 NTFANTLGAKEYLKNNILNSPERAKAKAFDNDPWNHSLPPRSTLQHPDYLOEYSTKYFYK 502
QY 206 -----AENPEYL 212
DB 503 QNGRIPIVAENPEYL 518

Search completed: April 28, 2003, 13:43:20
Job time: 8.35033 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 28, 2003, 13:37:49 ; Search time 14.515 Seconds
(without alignments)
3080.404 Million cell updates/sec

Title: US-09-821-883-25

Perfect score: 1182

Sequence: 1 GAGCGVHNRHRSSTRSGG.....STFKGTPAENPEYLGLDVP 217

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mmc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1026.5	86.8	1259	6 018735	018735 canis faml
2	961	81.3	367	11 08R2X1	08R2X1 mus musculu
3	906	76.6	165	4 014256	014256 homo sapien
4	894.5	75.7	412	4 08WYV0	08WYV0 homo sapien
5	165.5	14.0	1137	13 09W6F6	09W6F6 gallus galli
6	144.5	12.2	420	5 08S247	08S247 drosophila
7	144.5	12.2	446	5 09VZC2	09VZC2 drosophila
8	144	12.2	1209	11 09OX70	09OX70 rattus norv
9	140	11.8	585	12 041935	041935 muria herpe
10	137.5	11.6	358	4 008805	008805 homo sapien
11	136.5	11.5	438	10 039495	039495 cylindrothe
12	136	11.5	1210	11 09EP98	09EP98 mus musculu
13	135.5	11.5	297	4 016038	016038 homo sapien
14	134.5	11.4	379	13 091810	091810 xenopus lae
15	134	11.3	309	5 018751	018751 caenorhabd1
16	134	11.3	382	4 000599	000599 homo sapien

17	134	11.3	674	10 065672	065672 arabidopsis
18	133.5	11.3	440	4 08TE44	08TE44 homo sapien
19	133.5	11.3	530	11 008934	008934 mus musculu
20	132.5	11.2	322	10 009084	009084 lycopersico
21	132.5	11.2	442	10 039494	039494 cylindrothe
22	132.5	11.2	487	11 08VDA4	08VDA4 rattus norv
23	132.5	11.2	584	16 09FCJ3	09FCJ3 streptomyc
24	131.5	11.1	473	10 039620	039620 chlamydomon
25	131	11.1	903	4 09UPX1	09UPX1 homo sapien
26	130.5	11.0	1480	4 096C04	096C04 homo sapien
27	130.5	11.0	2135	4 043157	043157 homo sapien
28	130.5	11.0	2135	4 09UIV7	09UIV7 homo sapien
29	130	11.0	847	10 09XIB6	09XIB6 arabidopsis
30	130	11.0	3325	12 09IBT9	09IBT9 turkey herp
31	130	11.0	3342	12 09EBN3	09EBN3 turkey herp
32	129.5	11.0	476	5 09XTT6	09XTT6 caenorhabd1
33	129.5	11.0	1620	11 08V159	08V159 mus musculu
34	129	10.9	440	4 08TF74	08TF74 homo sapien
35	128	10.8	889	16 09F2N5	09F2N5 streptomyc
36	127.5	10.8	551	13 013003	013003 gallus galli
37	127.5	10.8	594	13 09DF69	09DF69 gallus galli
38	127.5	10.8	955	10 094F92	094F92 chlamydomon
39	127.5	10.8	3179	12 08V2A4	08V2A4 human herpe
40	127	10.7	369	10 040692	040692 oryza sativ
41	127	10.7	763	2 09XDH2	09XDH2 mycobacteri
42	126.5	10.7	356	12 09Q0B5	09Q0B5 herpes simp
43	126.5	10.7	734	11 088970	088970 mus musculu
44	126	10.7	562	16 0923X1	0923X1 rhizobium m
45	126	10.7	990	4 09UG03	09UG03 homo sapien

ALIGNMENTS

RESULT 1

018735 PRELIMINARY; PRT; 1259 AA.

ID 018735

AC 018735

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Erdb-2.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RM [1]

RP SEQUENCE FROM N.A.

RA Yokota H.;

RT "cDNA cloning of erdb-2 from canine mammary gland.";

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB008451; BAA23127.1; .

DR HSSP; P11362; IFC.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR000494; EGFR_L_domain.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR002174; Furin-like.

DR InterPro; IPR001245; Tyr_pkinase.

DR InterPro; IPR004019; YIP_motif.

DR Pfam; PF00757; Furin-like; 1.

DR Pfam; PF00069; Kinase; 1.

DR Pfam; PF01030; Recep_L_domain; 2.

DR Pfam; PF02757; YIP; 2.

DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SMO0261; Fu; 3.

DR SMART; SMO0219; Tyrc; 1.

DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

KW ATP-binding; Transferase; Tyrosine-protein kinase.

SO SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;

Query Match 86.8%; Score 1026.5; DB 6; Length 1259;
 Best Local Similarity 86.0%; Pred. No. 5.9e-69;
 Matches 191; Conservative 9; Mismatches 17; Indels 5; Gaps 1;

OY 1 GAGGWHRRSSSTRSGGDLTLGLEPSEEARSLASSEGASVDFGDLGMAK 60
 DB 1037 GAGGTARRHRSSSTRNGGELTLGLEPSEEPKSLASSEGASVDFGDLGMAK 1096
 OY 61 LQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRPQPSPRG 120
 DB 1097 LQSLPSDDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRPQPSPRG 1156
 OY 121 PAARPAAGATLER-----AKTLSPGKNGVAKDVFAGGAVENPEYLTPQGAAPQHP 175
 DB 1157 PPSRPAAGATLERPKTLSPKNGVAKDVFAGGAVENPEYLTPQGAAPQHP 1216
 OY 176 FSPAFLNLYWDDPPERGAPSTFKGTPTAENPEYLG 217
 DB 1217 FSPAFLNLYWDDPPERGAPSTFKGTPTAENPEYLG 1258

RESULT 2

OBR2X1 PRELIMINARY; PRT; 367 AA.
 AC OBR2X1;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical 40.2 kda protein.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC027080; AAH27080.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 367 AA; 40163 MW; 0BE0395F9E101B0 CRC64;

Query Match 81.3%; Score 961; DB 11; Length 367;
 Best Local Similarity 82.0%; Pred. No. 1.2e-64;
 Matches 178; Conservative 8; Mismatches 31; Indels 0; Gaps 0;

OY 1 GAGGWHRRSSSTRSGGDLTLGLEPSEEARSLASSEGASVDFGDLGMAK 60
 DB 150 GTGTAHRRSSSTRSGGDLTLGLEPSEEARSLASSEGASVDFGDLGMAK 209
 OY 61 LQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRPQPSPRG 120
 DB 210 LQSLPSDDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEYVQNPVRPQPSPRG 269
 OY 121 PAARPAAGATLERPKTLSPKNGVAKDVFAGGAVENPEYLTPQGAAPQHP 180
 DB 270 PPSRPAAGATLERPKTLSPKNGVAKDVFAGGAVENPEYLTPQGAAPQHP 329
 OY 181 DNLTYWDDPPERGAPSTFKGTPTAENPEYLG 217
 DB 330 DNLTYWDDPPERGAPSTFKGTPTAENPEYLG 366

RESULT 3

O14256 PRELIMINARY; PRT; 165 AA.
 AC O14256;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE C-erb B2/neu protein (Fragment).
 GN C-ERB B2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86070181; PubMed=2999974;
 RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J.,
 RA Seeburg P.H., Liberman T.A., Schlessinger J., Franke U.,
 RA Levinson A., Ullrich A.;
 RT "Tyrosine kinase receptor with extensive homology to EGF receptor
 RT shares chromosomal location with neu oncogene."
 RL Science 230:1132-1139(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94000386; PubMed=8104414;
 RA Sarkar F.H., Ball D.E., Li Y.W., Crisman J.D.;
 RT "Molecular cloning and sequencing of an intron of Her-2/neu (ERBB2)
 RT gene."
 RL DNA Cell Biol. 12:611-615(1993).
 DR EMBL: M95667; AAC37531.1; -
 FT NON_TER
 SQ SEQUENCE 165 AA; 17327 MW; A0C113BA308BFA6B CRC64;

Query Match 76.6%; Score 906; DB 4; Length 165;
 Best Local Similarity 100.0%; Pred. No. 6.7e-61;
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 AGSVDFGDLGMAKAGLQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEY 103
 DB 1 AGSVDFGDLGMAKAGLQSLPTDPSPLQRYSDPTVPLPSETDGYVAPLTCSPQPEY 60
 OY 104 NOPDVRPQPSPRGGLPAARPAAGATLERPKTLSPKNGVAKDVFAGGAVENPEYLTPQ 163
 DB 61 NOPDVRPQPSPRGGLPAARPAAGATLERPKTLSPKNGVAKDVFAGGAVENPEYLTPQ 120
 OY 164 GGAAPQHPPPAPFARDNLYWDDPPERGAPSTFKGTPTAEN 208
 DB 121 GGAAPQHPPPAPFARDNLYWDDPPERGAPSTFKGTPTAEN 165

RESULT 4

OBRVVO PRELIMINARY; PRT; 412 AA.
 AC OBRVVO;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical 44.7 kda protein.
 GN PP3659.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Huang Y., Zhou X.M., Zhang P.P., Jiang H.O., Qin W.X., Zhao X.T.,
 RA Wan D.F., Gu J.R.;
 RT "Novel human cDNA clones with function of inhibiting cancer cell
 RT growth."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF318349; AAL55856.1; -
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF02757; YLP_2.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 412 AA; 44702 MW; 034397FF3F27D2BC CRC64;

Query Match 75.7%; Score 894.5; DB 4; Length 412;
 Best Local Similarity 69.3%; Pred. No. 1.4e-59;
 Matches 181; Conservative 5; Mismatches 22; Indels 53; Gaps 4;

QY 1 GAGGAVHHRRSSSTRSGGDLTLEPSEEPAPRSLAPSGAGSDYFDGLGKAAG 60
 DB 150 GAGGAVHHRRSSSTRSGGDLTLEPSEEPAPRSLAPSGAGSDYFDGLGKAAG 209
 OY 61 LOSLTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEVYNODVROPSPREGPL 120
 DB 210 LOSLTHDPSPLQRYSEDPVPLPSETDGYVAPLTCSPQPEVYNODVROPSPREGPL 269
 OY 121 PAARPAGATLERAKTLPSPKNGVAVDFAFGAVENPEYLTPOGGAAPQ----- 170
 DB 270 PAARPAGATLERAKTLPSPKNGVAVDFAFGAVENPEYLTPOGGAALSPILLPSAQS 329
 OY 171 -----HPPPA---ESPAPDNLYWD-QDPPER----- 193
 DB 330 TTSITGRTHOSGGLHPPAPSKGLRQRTOSTWVWTCOCEPEGOVRRSPDVSSGREGULTS 389
 OY 194 -----GAPSTFKGTPAEN 208
 DB 390 AGIKRMEGPPTTSKGTCHARN 410

RESULT 5
 O9W6F6 PRELIMINARY; PRT: 1137 AA.

AC O9W6F6
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-JUN-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Receptor tyrosine kinase (Fragment).
 GN ERBB4
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HINDBRAIN;
 RX MEDLINE=99263203; PubMed=10328884;
 RA Dixon M., Lumsden A.;
 RT "Distribution of neurogranin-1 (nrg1) and erbB4 transcripts in
 embryonic chick hindbrain.";
 RL Mol. Cell. Neurosci. 13:237-258 (1999).
 DR HSSP; P11362; IFCK.
 DR InterPro: IPR000494; EGF_Ldomain.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR002174; Furin-like.
 DR InterPro: IPR001368; Tyr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF00757; Furin-like; 1.
 DR Pfam: PF00069; kinase; 1.
 DR Pfam: PF01030; Recep_Ldomain; 1.
 DR Pfam: PF02757; YLP; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR PRODOM; PD000001; Euk_kinase; 1.
 DR SMART; SM00261; FU; 3.
 DR SMART; SM00219; Tyrc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN1.
 KW Kinase; Tyrosine-protein kinase.
 FT NON TER
 SQ SEQUENCE 1137 AA; 127927 MW; 4D616436F87DC4F CRC64;

Query Match 14.0%; Score 165.5; DB 13; Length 1137;
 Best Local Similarity 26.7%; Pred. No. 0.00024;

Matches 62; Conservative 27; Mismatches 92; Indels 51; Gaps 10;

QY 16 RSGGDLTLEPSEEPAPR--PLAP--SEGAGSDVFDGLGKAAGLQSLPTDPSPL 72
 DB 899 RDGGYABQGV-PMYRAPGCTIPEAPVAGQATAFEDTCCNGTLRKQVATLAKEDST 957
 OY 73 QRYSDPTVPLPS-----ETDGYVAPLTCSPQPEVYNODVROPSPREGPLPA-R 124
 DB 958 QRYSDPTVPLPS-----ETDGYVAPLTCSPQPEVYNODVROPSPREGPLPA-R 1017
 OY 125 PAGATLERAKTLPSPKNGVAVDF-----AEGAVENPEYLTPOGGAAPQHPAPAF 176
 DB 1018 PEYH-----APNGQPAEDEVYNEPLYLNTFANTLENAEYL-----KNLPEKA 1062
 OY 177 SPAPDNLYWDQDPPERGA--PSTFKGTP-----AENPEYL 212
 DB 1063 KKAFDNPYWNHSLPPRSTLQHDYLOEXSTKYFKONGRIRPIVAENPEYL 1114

RESULT 6
 O8S247 PRELIMINARY; PRT: 420 AA.

AC O8S247
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RE17165P.
 GN CG15021.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyridae; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Paclet J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiliker S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY071124; AAL48746.1;
 SQ SEQUENCE 420 AA; 42947 MW; 96D62FFDC9F996E1 CXC64;

Query Match 12.2%; Score 144.5; DB 5; Length 420;
 Best Local Similarity 28.8%; Pred. No. 0.003;
 Matches 57; Conservative 11; Mismatches 79; Indels 51; Gaps 10;

QY 34 PRSPLAPSEAGSDVFDGLGKAAGL-----QSLPTDPSPLQRYSDPTVPLPS 85
 DB 37 PSVPF-PPPGSGNGIEDSGIGCPAPASAPSYGPPQTPRPPPPPO-----TPPAPR 90
 OY 86 ETDGYVAPLTCSPQPEVYNODVROPSPREGP--LPAARPAGATLERAKTLPSPKNGV 143
 DB 91 PS--YGPQTQPPRRPPQPTPSA-PAPPPSYGPPQTPRPPPPPOPTPSA----- 137
 OY 144 VKDVFAGAVENPEYLTPOGGAAPQHPAPAFSPADNLYWDQ-----PERGAPST 199
 DB 138 -----PAPPPSYGPPQ-----TPPAPPPQPTPSAPASVGPQPOPPAPAPSPSP 185
 OY 200 FKGTPTAENPEYLSIDVP 217
 DB 186 QPG-----PEYLPDP 197

RESULT 7
 O9VZC2 PRELIMINARY; PRT: 446 AA.

AC O9VZC2
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMBLER, 17, last annotation update)

DE CG15021 protein.

OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyridioidea; Drosophilidae; Drosophila.

OX NCBI_Taxid=7227;

RN [1]

RP SEQUENCE FROM 'N.A.

RC STRAIN=BERKELEY.

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,

RA Abdl J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,

RA Beeson K.Y., Berno P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadiot E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke K., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doul L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Haveston K.A., Helman T.J., Hernandez J.R., Houck C.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jaitai M., Kalush F., Kapen G.H., Ke Z., Krennson J.A., Kethum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Mekulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,

RA Palazzolo M., Plattan G.S., Pan S., Pollard J., Pui Y., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassenaar D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RL "The genome sequence of *Drosophila melanogaster*.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003481; AAF47902.1; -.

DR FMBase: FBgn0035544; CG15021.

DR InterPro: IPR003882; Pfam: extensin.

DR InterPro: IPR002965; Pfam: extensin.

DR PRINTS: PR01217; PRTCHEXTNSM.

DR PRINTS: PR01218; PSTTEXTNSM.

QO SEQUENCE 446 AA; 45728 MW; A2C9CE02A07F0542 CRC64;

	Query Match	12.2%	Score 144.5:	DB 5:	Length 446:	
	Best Local Similarity	28.8%:	Pred. NO. 0.0032:			
	Matches 57:	Conservative 11:	Mismatches 79:	Indels 51:	Gaps 10:	
OY	34	PRSLPASEGAGSDVFDGLGMGAAGL-----QSLPTDPSLQRSEDPYPLDS	85			
Dd	63	PSVFP-PPPGSNGLEDSGICGPAPASAPSYGPOTRPPPPPPO-----PIPPPAR	116			
OY	86	ETDGYAFLTCSPOPEYVNQDPVRPOPSPREGP--LPAPAGATLERAKTLSPKNCV	143			
Dd	117	PS-TGYPQOTOPRRPPQPRTPSA-PAPPSPSGPQTTPRRPPQRTPSA-----	163			
OY	144	VKDVAEFGGANENPYLLPGGAAQAQHPPAFSAFNLWYDDP-----PERCAPST	199			
Dd	164	-----PAAPPSTGGPO--TPPPRPPOPTSAPASYSAGPQQPAPQAPSPPSP	211			

```

QY      200  FKGTPTAENPEYLGLDVP  217
          |      ||| | |
Db      212  QPG-----PEYLPDPQP  223

```

RESULT 8	
Q9QX70	
ID Q9QX70	PRELIMINARY;
	PRT; 1209 AA

DT 01-MAY-2000 (TREMBLREl. 13, Created)

DT 01-JUN-2002 (Tremblay, 21, Last annotation update)

ECFR.

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OX NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.

RX MEDLINE=90258888; PubMed=2342466;

RA Earp H.S.;

RT encoded by an alternatively spliced transcript in normal rat tissue.";

[2]

RC STRAIN=FISHER: TISSUE=LIVER:

Submitted (NOV-1991) to the EMBL/GenBank/DBJ databases

SEQUENCE FROM N A
[2]
RN
BP

RC SIKAIN=FISHER; TISSUE=LIVER;
BA Guttridgc K Dawson T L Farr H S ;

DB EMBL: M37394: 24514008 1 : -
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases

DR HSSP; P11362; LEKG.
T4TCTPQ: TBP000404; ECFB

DR	InterPro; IPR000719; Euk_
DR	InterPro; IPR003174; Euk_

DR InterPro; IPR001245; Tyr-

DR Pfam; PF00069; pkinase; 1

DR PRINTS; TYRKINAS

DR SMART; SM00261; FU; 3.

DR PROSITE; PS00107; PROTEIN

DR PROSITE; PS00109; PROTEIN

SEQUENCE 1209 AA; 1348

Query Match	12.2
-------------	------

Matches 59: Conservative

47 DVFDGDI,GMGAARGT,OST,PT

1010 DVADEVT.TBOGFENSPS

77 EDGEMUND DEEST - DOVYADI

[illegible]

.....

● ●
——
——
● ●
● ●

● ●
● ●

——
——

QY 189 ----DP-----PERGAPSTFKGPTPAENPEYLGLDVP 217
 Db 1164 MSLDNPDYQODFFPKAKPMGICFKG-PTAENAEYLRAVAP 1202

RESULT 9

ID 041935 PRELIMINARY; PRT: 585 AA.
 AC 041935;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JUN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE Hypothetical 60.2 kDa protein.
 GN GAMMAHV. M6.
 OS murid herpesvirus 4.
 OC viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae.
 OX NCBI_TaxID=33708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WMMS;
 RX MEDLINE=97366649: PubMed-9223479;
 RA Virgin H.W. IV, Latreille P., Wamsley P., Hallsworth K., Weck K.E.,
 RA Dal Canto A.J., Speck S.H.;
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus
 RT 68."
 RL J. Virol. 71:5894-5904(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WMMS;
 RA Latreille P., Wamsley P., Waterston R.H.;
 RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U97553: AAB6392.1: -
 KW Hypothetical protein.
 SQ SEQUENCE 585 AA; 60160 MW; 85610AAB0C34827D CRC64;

Query Match 11.8%; Score 140; DB 12; Length 585;
 Best Local Similarity 26.5%; Pred. No. 0.0094;
 Matches 60; Conservative 8; Mismatches 90; Indels 68; Gaps 10;

QY 28 PSEEARSPRLAPSE-GAGSVYFPGDGLGMAKGLQSLPHDPPLQRTSDDPVLPSE 86
 Db 211 PSMGPDRPPRPPELPGSPTSPAPSRAGA-----RIPDL-PGLPSWGDPPRP 264
 QY 87 TDGVYALTCSPQPEYVNO-----PDVRPOPSPSRGP-----LPAARPGAT 129
 Db 265 ELGGSP--TSPAPSRAGARIPDLPGPLPSWGPDRPPRPPELPGSPTSPAPSRAGAR 322
 QY 130 LERAKTISPCKNGVYKDVAFGAVENPEYLTPOGAAPQ-PHPPAPFSPAFDN----- 182
 Db 323 IPDL-----PGP-----LPWGPDPDRPPRPPELPGSPTSPAPSR 358
 QY 183 -----LYYWDODPPERGAPSTFKGPTPAENPEYLGLDVP 217
 Db 359 AGARIPDLPGPLPSWGDPPRPPELPGSPTSPAPSRAGARIP 404

RESULT 10

ID 008805 PRELIMINARY; PRT: 358 AA.
 AC 008805;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PRBL protein (Fragment).
 GN PRBL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93304421: PubMed-8317492;

RA Azen E.A., Latreille P., Niece R.L.;
 RT "PRBL gene variants coding for length and null polymorphisms among
 RT human salivary Ps, Pm, PMS, and Pe proline-rich proteins (PRPs).";
 RL Am. J. Hum. Genet. 53:264-278(1993).
 DR EMBL: S62941: AAB27289.1: -
 FT NON-TER
 SQ SEQUENCE 358 AA; 35050 MW; DB7F878BD5EA759E CRC64;

Query Match 11.6%; Score 137.5; DB 4; Length 358;
 Best Local Similarity 26.6%; Pred. No. 0.0084;
 Matches 59; Conservative 17; Mismatches 75; Indels 71; Gaps 11;

QY 11 RSSSTRSGGDLTLGLPSEEARSPRLAP---SEG---AGSDVFDGLGMAKGLQSL 64
 Db 52 KSRSPRSPPK-PQGPFGGNGQPGPPPGKPGGPPGKNGKPGGPGGPGGPGGPGGPGG 106
 QY 65 PTHDPSPLQRYSEDPTVPLSEIDGYAPLTCSPQPEYVQPDVRRPP----- 113
 Db 107 ---PPQGDKSQSPRSP-PGKPGG-----PPQGGNQPGPPPPGKPGGPGGPGG 153
 QY 114 SPREGPLAARPGATLERAKTIS---PGKNGVYKDVAFGAVENPEYLTPOGAAPQ 169
 Db 154 NKPGPPPPGKPGGPPPGGKSGSPRSPGK-----PQGPFGGNGQPG 197
 QY 170 PHPPAPFSPAFDNLVYWDODPPERGA-----PSTFKGTP 204
 Db 198 GPPPPGKP-----QGPPGGGNRPGGPPPPGKPGGPP 230

RESULT 11

ID 039495 PRELIMINARY; PRT: 438 AA.
 AC 039495;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Alpha 2 frustulin.
 GN FRU ALPHA 2.
 OS Cyndrotheca frustiformis (Marine diatom).
 OC Eukaryota; stramenopiles; Bacillariophyta; Bacillariophyceae;
 OC Bacillariophycidae; Bacillariales; Bacillariaceae; Cyndrotheca.
 OX NCBI_TaxID=2853;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96314479: PubMed-8706728;
 RA Kroegeer N., Bergsdorf C., Sumper M.;
 RT "Frustulins: Domain conversation in a protein family associated with
 RT diatom cell walls."
 RL Eur. J. Biochem. 239:259-264(1996).
 DR EMBL: X99327: CAA67702.1: -
 SQ SEQUENCE 438 AA; 45947 MW; 3F894F503634B03 CRC64;

Query Match 11.5%; Score 136.5; DB 10; Length 438;
 Best Local Similarity 27.8%; Pred. No. 0.012;
 Matches 57; Conservative 20; Mismatches 55; Indels 73; Gaps 12;

QY 34 PRSLAPSEG-----AGSDV-----DGDLGMAKGLQSL----- 63
 Db 108 PGVPLEKCEGDCSDSCASSDLFCFPRNDVYVPGRGOSDSDSKDYCIKRTADAPGV 167
 QY 64 ---LPTHDPSPLOKRYSEDPTVPLSEIDGYAPLTCSPQPEYVQPDVRRPP---PSPRE 117
 Db 168 APVYPTADPSP-----DPT-PDPSPP-----TPPSPSPVNSPPDPPTPGPTPPTP 214
 QY 118 GPLAARPAAG-----ATLERAKTISG-----KNGVY---KDV-----FAFGAVE 155
 Db 215 GPTPASASGSDPPKATYFPLCLGCDGDDPDCEGDLICFORDANESVPGSGSSSD 274
 QY 156 --NPEYLTPOGAAPOPHPPAPSP 178
 Db 275 RSRDYCIKENPSTPPPTPPAPAP 299

```

RESULT 12
ID Q9EP98 PRELIMINARY: PRT: 1210 AA.
AC Q9EP98;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Epidermal growth factor receptor isoform 1.
GN EGFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/101, 129/SVJ, AND 129/SVEVTA;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Sincelar C.S., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Mahle N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/63;
RA Reiter J.L., Threadgill D.W., Eley G.D., Strunk K.E., Danielson A.J.,
RA Schell C., Pearsall R.S., Green P.J., Yee D., Lampland A.L.,
RA Balasubramaniam S., Crossley T.O., Magnuson T.R., James C.D.,
RA Mahle N.J.;
RT "Comparative genomic sequence analysis and isolation of human and
RT mouse alternative Egr transcripts encoding truncated receptor
RT isoforms."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF275366; AAG28045.1; -.
DR EMBL: AF275364; AAG28045.1; JOINED.
DR EMBL: AF275365; AAG28045.1; JOINED.
DR EMBL: AF275367; AAG24386.1; -.
DR HSSP: P11362; 1FGK.
DR MGD: MGI:95294; Egfr.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR000494; EGFR_L_domain.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00757; Furin-like; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF01030; Recep_L_domain; 2.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00261; Fv; 5.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
KM ATP-binding; Receptor; Transferase.
SQ SEQUENCE 1210 AA; 134840 MW; 62CD021C9DE32E18 CRC64;

Query Match 11.5%; Score 136; DB 11; Length 1210;
Best Local Similarity 26.4%; Pred. No. 0.043;
Matches 58; Conservative 23; Mismatches 63; Indels 76; Gaps 11;

```

```

OY 47 DVEPDGLGAKAAGLGLPHRDPSP-----LQRTS 76
DB 1011 DVAADADYLLPPOGFFNSPSTSRPLLSLSATSNSTVACINRNGSCRYKEDAFLORTS 1070
OY 77 EDPTVPLPSET--DGVAAPLTGSPQPEYVNPQVPRQPPSPRGCPPLAPAPACATLERAK 134
DB 1071 SDPTGAVTEQNDIDAF-----PVPEYVNO-SVPRKPAAGSVQNPVYHNQPLHP----- 1117

RESULT 13
ID Q16038 PRELIMINARY: PRT: 297 AA.
AC Q16038;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PRB1M protein (Fragment).
GN PRB1M.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93304421; PubMed=8317492;
RA Azen E.A., Latreille P., Niece R.L.;
RT "PRB1 gene variants coding for length and null polymorphisms among
RT human salivary Ps, PmF, PmS, and Pe proline-rich proteins (PRPs).";
RL Am. J. Hum. Genet. 53:264-278(1993).
DR EMBL: S62928; AAB27288.2; -.
FT NON-TER
SQ SEQUENCE 297 AA; 29046 MW; 1C7BE4CAB85B5F0 CRC64;

Query Match 11.5%; Score 135.5; DB 4; Length 297;
Best Local Similarity 25.5%; Pred. No. 0.0096;
Matches 64; Conservative 17; Mismatches 76; Indels 94; Gaps 13;

OY 11 RSSSTNSGGDLTLGLPSESEAPRSPLAD---SEG---AGSDVFDGDLGAKAGLQSL 64
DB 52 KSRSPRSPGK-PQGPPOGNGPOGPPPGKRGKPGGPPGKPGQ----- 106
OY 65 PTHDPSLQRYSDPTVPLPSELDGYVAPLTGSPQPEYVNPQVPRQPPS-----PR--- 116
DB 107 ----PPQGDKSQSPSP-rgkPQG-----PPQGGNQPGPPPPGKPGQPGQGG 153
OY 117 ---EGPLPAARPAATLERAKTSL---PGKNGVVKDVFAGAVENPEYLTLPQGAAPQ 169
DB 154 NRQGGPPPPGKPGQPPGQSKSPOSPPGK-----PQGPPOGNGNQ 197
OY 170 PHPPPAFSPAFDNLVYWDQDPPERG-----AIVSTFK 201
DB 198 GPPPPGPKP-----QGPPPOGNGKPGGPPPPGKPGQPPAOGSKSOSABAAPPKPG 248
OY 202 GTPTAE--NPE 210
DB 249 GPPQOEGGNPQ 259

RESULT 14
ID Q91810 PRELIMINARY: PRT: 379 AA.
AC Q91810;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Proline rich protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]

```

RP SEQUENCE FROM N.A.
 RC TISSUE=OOCYTE;
 RA Nishimatsu S.I., Satoshii, Oda, Naoto, Ueno;
 RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X68249; CAA48321.1; .
 DR InterPro: IPR003124; WH2; 1.
 DR Pfam: PF02205; WH2; 1.
 DR PRINTS: PRO1574; TUBBYPROTEIN.
 DR SMART: SM00246; WH2; 1.
 SO SEQUENCE 379 AA; 38859 MW; 84DC0FB24F971AAD CRC64;

Query Match 11.4%; Score 134.5; DB 13; Length 379;
 Best Local Similarity 28.0%; Pred. No. 0.015;
 Matches 59; Conservative 16; Mismatches 97; Indels 39; Gaps 9;

OY 2 AGGWNHHNRSSSTRSGGGDLTLGLEPSEEARPSRLAPSGE---AGSDVFDGDLGMAA 58
 DB 134 SGGRRHDDSDGSNRSPPVEGRARHPSLPDLSPSSSTSGMKHSSAPPPPPRRQA 193
 OY 59 KGLSLPTNHDSPLORYSEDPVPLPSETDGVAPLTCSPOPEYVQPDVRPQPPSP--- 115
 DB 194 -GAPAPSQNMKPYNR--EKPLPPTGHRAPAPVKKPPSPINSRPSAHSQPPPPPY 250
 OY 116 -----REGPLAARPPGATLERAKTISPCKNGYVKDVFAGAVENPE--YLTPOGGA 166
 DB 251 RQPTSLNGPSPSPINEPAPELPQRHNSLHRKTAGFVR-----GLAPPPQSVHLSP -CGN 304
 OY 167 APQHPPPAFSPAFNDLYWDDPPERGAP 197
 DB 305 RP---PPPA-----RDPGGAAP 320

RESULT 15

O18751 PRELIMINARY: PRT: 309 AA.
 AC O18751.
 DT 01-NOV-1996 (TRENBLREL. 01, Created)
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
 DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
 DE C50F7.5 protein.
 GN C50F7.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 CC Rhabditidae; Peleoderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Johnson, D., Steillyes L.;
 RT "The sequence of C. elegans cosmid C50F7.";
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U41557; AAA83301.1; .
 SO SEQUENCE 309 AA; 31085 MW; 6171574A991696D7 CRC64;

Query Match 11.3%; Score 134; DB 5; Length 309;
 Best Local Similarity 24.2%; Pred. No. 0.013;
 Matches 54; Conservative 25; Mismatches 74; Indels 70; Gaps 10;

OY 7 HHRHRSSTRSGGGDLTLGLEPSEEARPSRLAPSGEAGSDVFDGDLGMAAKG:QSLPT 66
 DB 31 HHHHKTAPRTSRGIATTTFAFTSSDL---PIAGSSSA-----
 OY 67 HDPSPLQRYSEDP-----TVPLPSE--TDGVAPLTCSPOPEYVQPDVRPQPPSPRGP 119
 DB 66 -----PVIASSADPILPITSVVPQPSNPSPGTVAP---SDESPSGPPS--PGVNPSEDP 116
 OY 120 LPAAPAGATLERAKTISPCKNGYVKDVFAGAVENPEYLTPOGGAPOP-----HPP 174
 DB 117 QPSGPPSPGPVDPSEDPQ-----SVEPSEDHQPSGPPSPGPVDPSEDPQ 162
 OY 175 AFSPAFNDLYWDDPPERGAPPSFTKGTPTAENPEYGLDVP 217
 DB 163 SVPSSED-----PQSGPPSPGPVDP--SEDPQPSGSSSP 195

Search completed: April 28, 2003, 13:41:41
 Job time : 20.515 secs

THIS PAGE BLANK (USPTO)